PCT WORLD INTELLECTUAL PR
Internation

INTERNATIONAL APPLICATION PUBLISHED UNI

WO 9602644A1

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, 16/18, A61K 38/17

A1

(11

(43) Internati nal Publicati n Date: 1 February 1996 (01.02.96)

(21) International Application Number: PCT/US95/08493

(22) International Filing Date: 6 July 1995 (06.07.95)

(30) Priority Data:

08/277,803	20 July 1994 (20.07.94)	US
08/384,710	1 February 1995 (01.02.95)	US

(81) Designated States: AU, CA, JP, MX, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: WOOD, Clive; 2 Hawthorne Place #17R, Boston, MA 02114 (US). CARUSO, Anthony; 29 Newbury Avenue, Methuen, MA 01844 (US).

(74) Agent: BROWN, Scott, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(54) Title: MLK RECEPTOR TYROSINE KINASES

(57) Abstract

Polynucleotides encoding novel receptor tyrosine kinases designated "mlk" are disclosed. mlk proteins and methods for their production, ligands for the mlk receptor and methods for their identification, and inhibitors of binding of mlk and its ligands and methods for their identification are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

MLK RECEPTOR TYROSINE KINASES

This application is a continuation-in-part of application Ser. No. 08/277,803, filed July 20, 1994.

Field of the Invention

The present invention relates to receptor tyrosine kinases, nucleic acids encoding such proteins, ligands thereto and methods of identifying inhibitors of activity of such proteins.

Background of the Invention

The receptors for almost all of the growth factors of biological interest belong to the tyrosine kinase or hematopoietin receptor families. Receptor tyrosine kinases (RTKs) include the receptors for M-CSF, kit ligand, neurotrophins and fibroblast growth factors. RTKs typically act as the first link in the chain of a responder cell which leads to biological effects of various growth factors. Ligand binding to the extracellular region of the RTK leads to activation of its intracellular kinase domain and autophosphorylation of tyrosine residues. This phosphorylation causes a cascade of other cellular events which ultimately result in exhibition of the activity caused by the ligand. See generally, Yarden et al., Ann. Rev. Biochem, 1988, 57, 443; Bolen et al., Advances in Cancer Res., 1991, 57, 103; Fantl et al., Annu. Rev. Biochem., 1993, 62, 453; Schlesinger et al., Neuron, 1992, 9, 383.

Because of their crucial position in the chain of events leading to the manifestation of biological activity, RTKs have recently been the subject of extensive study. Several RTKs have already been identified and, in many instances, cloned. Identified RTKs include: flt1

(deVries et al., Science, 1992, 255, 989); flt4 (Pajusola et al., Cancer Res., 1992, 52, 5738); flk1 and flk2 (U.S. Patent Nos. 5.270,458, 5,185,438; WO93/10136; WO92/17486; WO93/00349; Matthews et al., Cell, 1991, 65, 1143; Rosnet et al., Oncogene, 1991, 6, 1641); RYK (WO93/23429); KDR (WO92/14748); tie (WO93/14124); *Torpedo californica* RTK (Jennings et al., Proc. Natl. Acad. Sci. USA, 1993, 96, 2895); tyro10 (Lai et al., Oncogene, 1994, 9, 877); ptk3 (Sanchez et al., Proc. Natl. Acad. Sci. USA, 1991, 91, 1819); trkb (Middlemas et al., Mol. Cell Biol., 1991, 11, 143); and HEK (Wicks et al., Proc. Natl. Acad. Sci. USA, 1992, 89, 1611).

Some RTKs are differentially expressed in different tissues. For example, flk2 has been reported to be expressed in primitive hematopoietic cells, but not in mature hematopoietic cells. Thus, the differential expression of certain RTKs in a given tissue type may indicate a function for that RTK which is specific to such tissue. As a result, identification of tissue-specific RTKs may lead to the identification of RTKs involved in tissue-specific processes, and subsequently the ligands responsible for activating these tissue-specific processes. Therefore, it would be desirable to identify other tissue-specific RTKs to determine their relationship to tissue-specific functions.

Summary of the Invention

In accordance with the present invention, polynucleotides encoding novel RTKs are disclosed. The novel RTKs have been designated "muscle-localized kinase" or "mlk". "mlk" is used throughout the present specification to refer to both receptor proteins and polynucleotides encoding those proteins and to refer to proteins and polynucleotides from all

mammalian species.

In certain embodiments, the present invention provides for an isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of: (a) the nucleotide sequence of SEQ ID NO:1 from nucleotide 13 to nucleotide 1602; (b) the nucleotide sequence of SEQ ID NO:18 from nucleotide 1 to nucleotide 2580; (c) the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 2604; (d) a nucleotide sequence capable of hybridizing to a nucleic acid sequence specified in (a)-(c); (e) a nucleotide sequence varying from the sequence of a nucleotide sequence specified in (a)-(c) as a result of degeneracy of the genetic code; and (f) an allelic variant of a nucleotide sequence specified in (a)-(c). In particular embodiments, the nucleotide sequence encodes for protein having mlk receptor activity and/or mlk ligand binding activity. Polynucleotides encoding various mammalian mlk proteins (including murine and human mlk proteins) are included in the present invention. In other embodiments, the nucleotide sequence is operably linked to an expression control sequence.

The present invention also encompasses host cells transformed with the polynucleotide of the invention, including mammalian cells.

In other embodiments, the invention provides for a process for producing a mlk protein, said process comprising: (a) growing a culture of host cells transformed with a polynucleotide of the invention in a suitable culture medium; and (b) purifying the mlk protein from the culture. mlk protein produced by such methods is also provided by the present invention.

Further embodiments of the present invention provide isolated mlk protein comprising an amino acid sequence selected from the group consisting of: (a) the amino acid

sequence of SEQ ID NO:2; (b) the amino acid sequence of SEQ ID NO:2 from amino acids 1 to 156; (c) the amino acid sequence of SEQ ID NO:2 from amino acids 157 to 177; (d) the amino acid sequence of SEQ ID NO:2 from amino acids 178 to 530; (e) the amino acid sequence of SEQ ID NO:2 from amino acids 242 to 517; (f) the amino acid sequence of SEQ ID NO:15; (g) the amino acid sequence of SEQ ID NO:15 from amino acids 22 to 478; (h) the amino acid sequence of SEQ ID NO:19; (i) the amino acid sequence of SEQ ID NO:19 from amino acids 22 to 486; (j) the amino acid sequence of SEQ ID NO:19 from amino acids 487 to 507; (k) the amino acid sequence of SEQ ID NO:19 from amino acids 508 to 860; (l) the amino acid sequence of SEQ ID NO:19 from amino acids 572 to 847; (m) the amino acid sequence of SEQ ID NO:21; (n) the amino acid sequence of SEQ ID NO:21 from amino acids 22 to 494; (o) the amino acid sequence of SEQ ID NO:21 from amino acids 495 to 515; (p) the amino acid sequence of SEQ ID NO:21 from amino acids 516 to 868; (q) the amino acid sequence of SEQ ID NO:21 from amino acids 580 to 855; (r) fragments of (a)-(q) having mlk receptor activity; and (s) fragments of (a)-(q) having mlk ligand binding activity. Pharmaceutical compositions comprising mlk protein and a pharmaceutically acceptable carrier are also contemplated by the present invention, as are compositions comprising antibodies (polyclonal and monoclonal) which specifically react with mlk protein. Polynucleotides encoding each of these proteins are also provided by the present invention.

The present invention also encompasses methods of identifying a mlk receptor ligand, said method comprising: (a) providing a sample containing a potential source of mlk ligand; (b) contacting said sample with a protein having mlk receptor activity or mlk ligand binding activity; and (c) collecting materials binding to the protein. mlk receptor ligands

NOT TO BE TAKEN INTO CONSIDERATION FOR THE PURPOSES
OF INTERNATIONAL PROCESSING

NOT TO BE TAKEN INTO CONSIDERATION FOR THE PURPOSES
OF INTERNATIONAL PROCESSING

These corresponding polynucleotide fragments are also part of the present invention. Modified polynucleotides as described above may be made by standard molecular biology techniques. including site-directed mutagenesis methods which are known in the art or by the polymerase chain reaction using appropriate oligonucleotide primers.

For the purposes of the present invention, a protein has "mlk receptor activity" when binding of a ligand to such protein results in tyrosine phosphorylation as measured by the tyrosine phosphorylation assay described below or other suitable assays. A protein has "mlk ligand binding activity" when it binds mlk ligand as measured by the ligand binding assay described below or other suitable assays.

mlk protein or fragments thereof having mlk ligand binding activity may be fused to carrier molecules such as immunoglobulins. For example, soluble forms of the mlk protein such as those fragments comprising amino acids 1-125 of SEQ ID NO:2 may be fused through "linker" sequences to the Fc portion of an immunoglobulin.

The invention also encompasses allelic variations of the nucleotide sequence as set forth in SEQ ID NO:1, that is, naturally-occurring alternative forms of the isolated polynucleotide of SEQ ID NO: 1 which also encode proteins having mlk receptor activity or mlk ligand binding activity. Also included in the invention are isolated polynucleotides which hybridize to the nucleotide sequence set forth in SEQ ID NO:1 under highly stringent (0.2xSSC at 65°C), stringent (e.g. 4xSSC at 65°C or 50% formamide and 4xSSC at 42°C), or relaxed (4xSSC at 50°C or 30-40% formamide at 42°C) conditions. Isolated polynucleotides which encode mlk protein but which differ from the nucleotide sequence set forth in SEQ ID NO:1 by virtue of the degeneracy of the genetic code are also encompassed by the present invention.

Variations in the nucleotide sequence as set forth in SEQ ID NO:1 which are caused by point mutations or by induced modifications which enhance mlk receptor activity or mlk ligand binding activity, half-life or production level are also included in the invention.

The isolated polynucleotides of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the mlk protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means enzymatically or chemically ligated to form a covalent bond between the isolated polynucleotide of the invention and the expression control sequence, in such a way that the mlk protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the mlk protein. Any cell type capable of expressing functional mlk protein may be used. Suitable mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK, Rat2, BaF3, 32D, FDCP-1, PC12 or C2C12 cells.

The mlk protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression

vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. Soluble forms of the mlk protein may also be produced in insect cells using appropriate isolated polynucleotides as described above.

Alternatively, the mlk protein may be produced in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium,* or any bacterial strain capable of expressing heterologous proteins.

The mlk protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a polynucleotide sequence encoding the mlk protein.

The mlk protein of the invention may be prepared by growing a culture transformed host cells under culture conditions necessary to express the desired protein. The resulting expressed protein may then be purified from the culture medium or cell extracts. Soluble forms of the mlk protein of the invention can be purified from conditioned media. Membrane-bound forms of mlk protein of the invention can be purified by preparing a total membrane fraction from the expressing cell and extracting the membranes with a non-ionic

detergent such as Triton X-100.

The mlk protein can be purified using methods known to those skilled in the art. For example, the mlk protein of the invention can be concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. Following the concentration step, the concentrate can be applied to a purification matrix such as a gel filtration medium. Alternatively, an anion exchange resin can be employed, for example, a matrix or substrate having pendant diethylaminoethyl (DEAE) groups. The matrices can be acrylamide, agarose, dextran, cellulose or other types commonly employed in protein purification. Alternatively, a cation exchange step can be employed. Suitable cation exchangers include various insoluble matrices comprising sulfopropyl or carboxymethyl groups. Sulfopropyl groups are preferred (e.g., S-Sepharose® columns). The purification of the mlk protein from culture supernatant may also include one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; or by hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or by immunoaffinity chromatography. Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the mlk protein. Some or all of the foregoing purification steps, in various combinations or with other known methods, can also be employed to provide a substantially purified isolated recombinant protein.

Preferably, the isolated mlk protein is purified so that it is substantially free of other mammalian proteins.

The mlk proteins of the present invention may also be used to identify and isolate mlk binding ligands. For example, mlk proteins having mlk ligand binding activity may be anchored to a solid support (such as in a microtiter plate or a binding affinity column) and exposed to possible sources of mlk ligands. Species binding to the mlk protein can then be eluted from the mlk protein and collected for further characterization. Sources of mlk binding ligands include among others blood, cerebrospinal fluid, extracted neuronal, muscular and splenic tissues, and membranes and conditioned media from cell lines derived from neuronal, muscular and hematopoietic tissue. mlk binding ligands may also be used in pharmaceutical compositions as described above with respect to the mlk protein. Ligands may also be identified in this manner which bind to the intracellular region of the mlk protein.

The invention also includes ligands which bind to the mlk receptor ("mlk receptor ligands"). The ligands may be growth factors that occur naturally in a mammal (either the same mammal or a different mammal from which the mlk receptor to which the ligand binds was isolated). The ligand may be isolated and purified, or be present on the surface of ligand-expressing cell populations. The ligand may also be a molecule that does not occur naturally in a mammal. The ligand may also be a non-protein molecule that acts as a ligand when it binds to or otherwise come into contact with, a mlk receptor protein. In cases where a ligand is a protein, polynucleotides encoding such ligands are also within the scope of the present invention. Once a protein ligand has been identified, the ligand-encoding polynucleotides may be obtained by expression cloning (Wong, Genetic Engineering, ed. by J.K. Setlow, 1990, Plenum Press, New York, vol. 12, p. 297) or by synthetic means.

The mlk proteins of the invention may also be used to screen for agents which are

capable of binding to mlk protein (either the extracellular or intracellular domains) and thus may act as inhibitors of normal ligand binding. Binding assays using a desired binding protein, immobilized or not, are well known in the art and may be used for this purpose using the mlk protein of the invention. Purified cell based or protein based (cell free) screening assays may be used to identify such agents. For example, mlk protein may be immobilized in purified form on a carrier and binding to purified mlk protein may be measured in the presence and in the absence of potential inhibiting agents. A suitable binding assay may alternatively employ purified mlk protein immobilized on a carrier, with a soluble form of mlk protein of the invention. Any mlk protein exhibiting mlk ligand binding activity may be used in the screening assays described above.

In such a screening assay, a first binding mixture is formed by combining a mlk binding ligand and mlk protein, and the amount of binding in the first binding mixture (B_o) is measured. A second binding mixture is also formed by combining a mlk binding ligand, mlk protein, and the compound or agent to be screened, and the amount of binding in the second binding mixture (B) is measured. The amounts of binding in the first and second binding mixtures are compared, for example, by performing a B/B_o calculation. A compound or agent is considered to be capable of inhibiting binding if a decrease in binding in the second binding mixture as compared to the first binding mixture is observed. The formulation and optimization of binding mixtures is within the level of skill in the art, such binding mixtures may also contain buffers and salts necessary to enhance or to optimize binding, and additional control assays may be included in the screening assay of the invention.

Compounds found to reduce the binding activity of mlk protein to ligand to any

degree, preferably by at least about 10%, more preferably greater than about 50% or more, may thus be identified and then secondarily screened in other binding assays and in vivo assays. By these means compounds having inhibitory activity for mlk binding which may be suitable as therapeutic agents may be identified.

It is believed that the mlk receptor plays an important role in skeletal muscle function. As described below, mlk expression is localized in skeletal muscle, spleen and lung (see Northern blot reproduced in Fig. 3). Furthermore, murine mlk maps to the proximal arm of murine chromosome No. 4. A mutant murine strain mapping in this proximity, "vacillans" (vc), is characterized by half normal muscle mass, gait problems, lack of muscle coordination, slowed behavior and reflexes, metabolic abnormalities, mental deficiencies and drastic reduction of body fat (see, Sirlin, J. Genetics, 1956, 54, 42). By synteny, it is believed that human mlk will map to the p or q arm of human chromosome 9. It has been reported that a condition denominated Fukuyama congenital muscular dystrophy ("FCMD") maps to human chromosome 9q31-33 (Toda et al., Nature Genetics, 1993, 5, 283), the expected locus of human mlk. FCMD is characterized by involvement of facial muscle and the central nervous system, brain malformation, progressive muscular dystrophy, mental retardation, autosomally recessive inheritance and joint contractures (Osawa et al., Acta Paediatr. Jpn., 1991, 33, 261). Familial dysautonomia ("DYS"), or Riley-Day syndrome, has also been mapped to human chromosome 9q31-33 (Blumenfeld et al., Nature Genetics, 1993, 4, 160). This hereditary disease affects the development and survival of sensory, sympathetic and some parasympathetic neurons. DYS is characterized by neurological deficiency, with a range of other features including skeletal muscle abnormality. The similarity between the mlk receptor and the *Torpedo* RTK also points toward

muscular involvement, since the *Torpedo* RTK was isolated from the electric organ of electric rays which is homologous to muscle tissue. Each of these findings implicates mlk receptor in muscle biology, perhaps playing a role in synapse formation and function.

As a result, isolated mlk proteins and mlk receptor ligands may be useful in treatment of various medical conditions in which the mlk receptor is implicated or which are effected by the activity (or lack thereof) of the mlk receptor (collectively "mlk-related conditions"). Mlk-related conditions include without limitation muscle-related disorders, diseases of the nervous system (including infections), vascular disorders, trauma, metabolic derangements, demyelinating diseases (including multiple schlerosis), neuronal diseases (including Alzheimer's disease, Parkinson's disease and Huntington's chorea; and including motor neuron diseases such as amyotrophic lateral sclerosis, primary lateral sclerosis and Werdnig-Hoffmann disease), epilepsy, syringomyelia, peripheral neuropathy, congenital anomalies and tumors. Muscle-related conditons for treatment include without limitation muscular dystrophies (such as severe and benign X-linked muscular dystrophy, limb-girdle dystrophy, facioscapulohumeral dystrohpy, myotonic dystrophy, distal muscular dystrophy, progressive dystrophic ophthalmoplegia, oculopharyngeal dystrophy and FCMD), DYS, congenital myopathy, myotonia congenita, familial periodic paralysis, paroxysmal myoglobinuria, myasthenia gravis, Eaton-Lambert syndrome, secondary myasthenia, and denervation atrophy. Given the importance of mlk RTK in muscle, perhaps especially in the establishment and function of synapses, it is also believed that the mlk protein and ligand will be of value in tissue repair.

It is also believed that the mlk receptor may play a role in the development and

growth of hematopoietic cells. As a result, mlk proteins and mlk receptor ligands may be useful in treatment of suppressed hematopoiesis, e.g., as a result of chemotherapy or associated with bone marrow transplantation. In addition, they may be useful for *in vitro* and *in vivo* stimulation of hematopoietic cells prior to transplantation, perhaps in peripheral stem cell transplantation. Furthermore, they may act to modulate the differentiation and maturation of hematopoietic cells, including lymphocytes. Thus, they may be of value, for example, in stimulation of antigen-specific cytotoxic T-cell action.

In situ hybridization experiments have also detected expression of mlk in fetal bone. As a result, mlk proteins or mlk receptor ligands may also be useful for treatment of various bone-related disorders, including without limitation those associated with bone loss (including that associated with osteoporosis, post-menopausal osteoporosis, senile osteoporosis, idiopathic osteoporosis, Pagets disease, multipe myeloma, and hypogonadal conditions). mlk proteins or mlk receptor ligands may also be useful as stimulators or inhibitors of osteoclasts and other cells involved in bone metabolism.

The mlk proteins or mlk receptor ligands may also be used to promote the growth of bone and cartilage and, therefore, have application in the healing of bone fractures and cartilage defects in humans and other animals. Such agents may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by such an agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery. These agents may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate

growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A variety of osteogenic, cartilage-inducing and bone inducing factors have been described. See. e.g., European patent applications 148,155 and 169,016 for discussions thereof.

The proteins of the invention may also be used in wound healing and related tissue repair. The types of wounds include, but are not limited to burns, incisions and ulcers. (See. e.g. PCT Publication WO84/01106 for discussion of wound healing and related tissue repair).

Isolated mlk proteins and mlk receptor ligands, purified from cells or recombinantly produced, may be used as a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may contain, in addition to mlk protein or ligand and carrier, diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-15, G-CSF, stem cell factor, and erythropoietin. The pharmaceutical composition may contain thrombolytic or anti-thrombotic factors such as plasminogen activator and Factor VIII. The pharmaceutical composition may further contain other anti-inflammatory agents. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with isolated mlk protein or ligand, or to minimize side effects caused by the isolated mlk protein or ligand. Conversely, isolated mlk protein may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic

or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

The pharmaceutical composition of the invention may be in the form of a liposome in which isolated mlk protein or ligand is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers which in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, e.g., amelioration of symptoms of, healing of, or increase in rate of healing of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of isolated mlk protein or ligand is administered to a mammal. Isolated mlk protein or ligand may be administered in accordance with the method of the

invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, isolated mlk protein or ligand may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering isolated mlk protein or ligand in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of isolated mlk protein or ligand used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, or cutaneous, subcutaneous, or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of isolated mlk protein or ligand is administered orally, isolated mlk protein or ligand will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% isolated mlk protein or ligand, and preferably from about 25 to 90% isolated mlk protein or ligand. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol.

When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of isolated mlk protein or ligand, and preferably from about 1 to 50% isolated mlk protein or ligand.

When a therapeutically effective amount of isolated mlk protein or ligand is administered by intravenous, cutaneous or subcutaneous injection, isolated mlk protein or ligand will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to isolated mlk protein an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additive known to those of skill in the art.

The amount of isolated mlk protein or ligand in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of isolated mlk protein or ligand with which to treat each individual patient. Initially, the attending physician will administer low doses of isolated mlk protein or ligand and observe the patient's response. Larger doses of isolated mlk protein or ligand may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not generally increased further. It is contemplated that the various

pharmaceutical compositions used to practice the method of the present invention should contain about 0.1 μg to about 100 mg of isolated mlk protein or ligand per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the isolated mlk protein or ligand will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Isolated mlk protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the mlk protein and which may inhibit ligand binding to the mlk receptor. Such antibodies may be obtained using the entire mlk protein as an immunogen, or by using fragments of mlk protein such as the soluble mature mlk protein. Smaller fragments of the mlk protein may also be used to immunize animals. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Additional peptide immunogens may be generated by replacing tyrosine residues with sulfated tyrosine residues. Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J.Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, et al., FEBS Lett. 211, 10 (1987).

Neutralizing or non-neutralizing antibodies (preferably monoclonal antibodies) binding to mlk protein or mlk ligand may also be useful therapeutics for certain tumors and also

in the treatment of conditions described above. These neutralizing monoclonal antibodies are capable of blocking the ligand binding to the mlk protein.

Example 1

Isolation of Murine mlk cDNA

A fragment of the murine mlk cDNA was isolated using PCR amplification of first strand cDNA with primers specific to conserved motifs of receptor tyrosine kinases. Bone marrow from C57BL/6J mice treated two days previously with 150 mg/kg 5-fluorouracil was depleted of B220+, Gr1+, CD3+ and Mac1+ cells, and then the cells bearing the Sca+ antigen were isolated by fluorescence-activated cell sorting (FACS). Total cellular RNA was purified from the sorted cells using the acid phenol technique.

First strand cDNA was synthesized using random hexamer primers and Moloney murine leukemia virus reverse transcriptase (Perkin Elmer GeneAmp RNA PCR reagents). The PCR scheme used to isolate the murine cDNA is summarized in Fig. 2. The cDNA was amplified using the primers of SEQ ID NO:3 and SEQ ID NO:4. The SEQ ID NO:3 and SEQ ID NO:4 primers were directed to the IHRDL and SDVWS motifs, respectively. The reaction was carried out in a Perkin Elmer 9600 Thermal Cycler, using the following conditions: 94°C for 30 sec., 37°C for 15 sec., 72°C for 15 sec. (5 cycles): 94°C for 30 sec., 50°C for 15 sec., 72°C for 15 sec. (25 cycles).

The product of this amplification was purified by Centricon 30 microcentrifugation (Amicon), and then amplified again using the primers of SEQ ID NO:5 and SEQ ID NO:6, directed to the DLAAR and WLALES motifs, respectively. The reaction was carried out under

the conditions described above.

The second primers contained EcoRI sites at their 5' ends to facilitate sub-cloning. Thus, the products of the second reaction were digested with EcoRI, and examined by gel electrophoresis using low melting temperature agarose and ethidium bromide staining. the approximately 150 bp band expected was excised and eluted. The isolated fragments were ligated into EcoRI digested pBluescript KSII+ (Stratagene), and used to transform competent InvαF' E. coli (Invitrogen). Transformants were selected on L-amp plates containing 100 μg/ml ampicillin. and then randomly selected for miniplasmid preparations and analysis by nucleotide sequencing.

The motifs of the first PCR are highly conserved in protein tyrosine kinases (Hanks and Quinn, Methods in Enzymol., 1991, 200, 38) and this amplification selects for a population of cDNAs derived from cytosolic and receptor tyrosine kinases. The motifs of the second PCR are specific to RTKs. Analysis of products of the second PCR amplification showed this strategy to be successful. No cytosolic tyrosine kinases were obtained. However, previously described RTKs, such as flk-1, flk-2 and c-kit, were found. The sequence of one transformant was found to be novel and not previously described.

Longer cDNA clones of this novel sequence. including pred2#2 (SEQ ID NO:1), were obtained from a murine embryo (day 15) cDNA library in the lambda SHlox vector obtained from Novagen (Madison, WI). This library was hybridized with an oligonucleotide probe derived from the novel transformant (SEQ ID NO:7), which was radiolabelled with gamma-^{32}P-ATP. pred2#2 was deposited as an EcoRI/HindIII insert in pSHlox (Novagen, Wisconsin) with ATCC on July 20. 1994 as accession number 69659.

Example 2

Localization of mlk Expression

Expression of mlk mRNA is examined by Northern blot hybridization analysis of RNA isolated from primary tissues and cell lines. Cellular RNA is isolated by conventional techniques such as acid phenol isolation, guanidinium thiocyanate denaturation/density gradient centrifugation in cesium chloride or RNA Stat-60 isolation (Tel-Test Inc., TX). The mRNA fraction is selected on the basis of binding to oligo dT_{12-18}, using either oligo dT-cellulose chromatography, or binding to oligo dT-biotin and then magnetic separation after binding to streptavidin-paramagnetic particles (Promega, Inc., WI). RNA is exposed to formaldehyde-agarose gel electrophoresis and then transferred to nitrocellulose or nylon membranes by capillary blotting. Northern blot filters of murine and human tissue RNAs can also be purchased from Clontech, CA.

The first northern blot analysis of murine tissue RNAs was carried out with a murine mlk cDNA probe labelled with α-^{32}P-dCTP by random hexamer priming and the Klenow fragment of *E. coli* DNA polymerase I. Two mRNA species were found to hybridize to this probe, of approximately 7.0 and 4.0 kb. Both mlk mRNAs were detected in spleen and lung poly A^+ RNA, and at higher levels in skeletal muscle poly A^+ RNA. Subsequent northern blot experiments have not detected expression of mlk mRNAs in spleen or lung tissue and have confirmed high levels of mlk expression in skeletal muscle.

Analysis of mlk mRNA expression by northern blot in murine embryo polyA$^+$ RNA revealed that mlk is expressed at least as early as day 7 after fertilization.

Expression of mlk mRNA is also examined in tissue sections or whole mount

embryos by *in situ* hybridization with radiolabelled or chemically modified RNA probes. Antisera against mlk protein are generated by immunization of rabbits, mice, hamsters, goats or sheep with synthetic mlk peptides, such as the C-terminal 12-20 residue peptide conjugated to carrier. Antibodies against the extracellular portion of mlk are obtained by immunization with recombinant fusion proteins such as mlk-thioredoxin, mlk-maltose-binding protein, mlk-glutathione-S-transferase. These antibodies are used to localize mlk protein by immunohistochemistry. Antibodies against the extracellular region of mlk are used to isolate mlk expressing cells, *e.g.*, from bone marrow or muscle, by fluorescence-activated cell sorting, in conjunction with a secondary antibody that recognizes the primary anti-mlk, and is conjugated to a fluorochrome such as phycoerythrin or fluorescein isothiocyanate. Using additional antibodies conjugated to fluorochromes against known cell surface markers in conjunction with the anti-mlk protein will permit a phenotyping of the subsets of cells expressing mlk.

The expression of mlk polypeptide is also examined by immunoprecipitation and *in vitro* kinase assay. Cell or tissue extracts are prepared in 1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 150 mM sodium chloride, 20 mM Tris-HCl pH 8.0, or alternatively in 1% (v/v) Nonidet P-40, 1% (w/v) sodium deoxycholate, 0.1% (w/v) SDS, 25 mM Tris-HCl pH 8.0, 150 mM sodium chloride. Both solutions are supplemented with 1 mM sodium vanadate, 1 μg/ml leupeptin, 1 mM Pefabloc SC, 1 μg/ml aprotinin. Antisera raised against the C-terminal mlk peptide and protein A-Sepharose are then used to precipitate the mlk polypeptide. The immunoprecipitate is washed three times in the above extraction buffer, then three times in kinase assay buffer (20 mM HEPES, pH 7.3, 10 mM manganese chloride, 5 mM magnesium chloride) and then resuspended in 50 μl kinase assay buffer containing 2 μl of γ-^{32}P-ATP (10

μCi/μl; 1,000 Ci/mmol). After incubation for 15 minutes, the immunoprecipitate is washed three times in 20 mM HEPES, pH 7.3, 5 mM ethylene glycol-bis(β-amino-ethyl-ether) N,N,N',N'-tetraacetic acid (EGTA), 1mM rATP, and then boiled for 5 minutes in 1% SDS, 100 mM Tris-HCl pH 6.8, 50 μl/ml β-mercaptoethanol, prior to SDS-PAGE electrophoresis and autoradiography..

Example 3

Chromosomal Mapping of Murine mlk

Murine mlk was chromosomally mapped by single-strand conformation polymorphism analysis (Beier, Mammalian Genome, 1993, 4, 627) using the following oligonucleotides:

1F: CAGACTGTGAGCTGGAGGAAC (SEQ ID NO:8)

1R: GGATTATGTAGGAACGTAACC (SEQ ID NO:9)

2F: CGAAATAGGTTGGAGATACAGG (SEQ ID NO:10)

2R: CGTGTCTTGACTGGTGGAAGAAGG (SEQ ID NO:11).

The murine mlk gene was mapped genetically by SSCP analysis in the BxD recombinant inbred (R1) series. The murine mlk gene was mapped to the proximal arm of murine chromosome 4. By synteny, this region is predicted to correspond to human chromosome 9, p or q.

Example 4

Assay for mlk Receptor Activity

Cells expressing mlk protein are used in tyrosine phosphorylation assays for mlk receptor activity. NIH 3T3 or Rat2 cells are stably transfected with mlk expression plasmids. For 12-24 hours prior to the assay, the cells are maintained under standard growth conditions, except that only 0.5% fetal bovine serum is used in the culture medium. Monolayers of these cells, or suspensions prepared by detachment of cells by brief treatment with 20 mM EGTA in PBS, are stimulated for 5 min. at 37°C with ligand or antibodies directed to the extracellular region. Cells are then pelleted, and solubilized in either ice-cold 1% (v/v) Nonidet P-40, 1% (w/v) sodium deoxycholate, 0.1% (w/v) SDS, 25 mM Tris-HCl pH 8.0, 150 mM sodium chloride, or ice-cold 1% (v/v) Nonidet P-40, 10% (v/v) glycerol, 20 mM Tris-HCl pH 8.0, 150 mM sodium chloride. Both extraction buffers are supplemented with 1 mM sodium vanadate, 1 μg/ml leupeptin, 1 μg/ml aprotinin, 1 mM Pefabloc SC.

The solubilized proteins are then immunoprecipitated with rabbit antibodies directed against the mlk protein, such as the C-terminal peptide epitope, followed by protein A-sepharose. The immunoprecipitated material is then exposed to reducing SDS-PAGE electrophoresis and transferred to a nitrocellulose filter by Western blotting.

The amount of phosphotyrosine in the immunoprecipitated mlk protein is then assayed by incubation of the filter with a monoclonal antibody specific for phosphotyrosine (e.g., the 4G10 monoclonal antibody available from Upstate Biotechnology, New York). The antibody is either directly conjugated to horseradish peroxidase, or used as a biotin conjugate, followed by incubation with streptavidin-horseradish peroxidase. Binding is visualized with enhanced chemiluminescence reagents (commercially available from Amersham, Illinois) and exposure to x-ray film.

An increase in phosphotyrosine in samples treated (e.g., with a potential ligand source) relative to samples treated only with fresh medium, are taken to indicate induction of receptor activity. Positive samples can be further studied by examining the time course of phosphotyrosine induction on the mlk polypeptide.

Induction of phosphotyrosine on mlk protein can also be examined by immunoprecipitation with anti-phosphotyrosine conjugated to agarose, followed by Western blotting with anti-phosphotyrosine, or Western blotting of whole cell extracts with anti-phosphotyrosine.

Techniques for the study of phosphotyrosine on proteins have been described (Cooper et al., Methods in Enzymol., 1983, 99, 387).

Example 5

Assay for mlk Ligand Binding Activity

Binding of the mlk ligand to the extracellular region of the mlk receptor protein will specifically cause a rapid induction of phosphotyrosine on the receptor protein. Assays for ligand binding activity as measured by induction of phosphorylation is described in Example 4.

Activation of the receptor by ligand binding can also be determined by measuring cell growth or differentiation. For example, NIH 3T3 or Rat2 fibroblasts are stably transfected with expression vectors directing synthesis of mlk protein. A mitogenic response in these transfected cells as a result of ligand binding is measured by examining their growth in semi-solid medium, as described for other RTKs expressed in transfected NIH 3T3 or Rat2 cells (Cordon-Cardo et al., Cell. 1991, 66173; Glass et al., Cell. 1991, 66. 405; Dosil et al., Mol.

Cell Biol., 1993, 13, 6572; Maroc et al., Oncogene, 1993, 8, 909). In the presence of binding ligand, these cells will acquire the ability to grow in semi-solid media.

Induction of growth can also be measured by examining c-fos induction. The 3T3 or Rat2 cells are placed in 0.5% fetal bovine serum in DMEM for 24 hours prior to ligand addition. Addition of ligand or serum (positive control) will cause an induction of expression of the c-fos immediate response gene. Samples are harvested at zero time (prior to ligand addition), 30 min., 1 hour, 2 hours and 3 hours after ligand or serum addition. Cells harvested are used to isolate total cellular RNA, which is then examined by Northern blot analysis with a c-fos cDNA probe.

Alternatively, a soluble form of the mlk extracellular region is produced and used to detect ligand binding. A DNA construct is prepared in which the extracellular region encoding the mlk cDNA (truncated prior, preferably immediately prior, to the predicted transmembrane domain) is ligated in frame to a cDNA encoding the hinge C_H2 and C_H3 domains of a human immunoglobulin (Ig) $\gamma 1$. This construct is generated in an appropriate expression vector for COS cells, such as pEDΔC or pMT2. The plasmid is transiently transfected into COS cells. The secreted mlk-Ig fusion protein is collected in the conditioned medium and purified by protein A chromatography.

The purified mlk-Ig fusion protein is used to demonstrate ligand binding in a number of applications. The soluble ligand can be coated onto the surface of an enzyme-linked immunosorbent assay (ELISA) plate, and then additional binding sites blocked with bovine serum albumin or casein using standard ELISA buffers. The mlk-Ig fusion protein is then bound to the solid-phase ligand, and binding is detected with a secondary goat anti-human Ig conjugated

to horseradish peroxidase. The activity of specifically bound enzyme can be measured with a colorimetric substrate, such as tetramethyl benzidine and absorbance readings.

The mlk ligand may also be expressed on the surface of cells, for example by possessing a transmembrane domain or glucosyl phosphatidyl inositol (GPI) linkage. Cells expressing the membrane bound ligand can be identified using the mlk-Ig fusion protein. The soluble mlk-Ig fusion is bound to the surface of these cells and detected with goat anti-human Ig conjugated to a fluorochrome, such as fluorescein isothiocyanate and flow cytometry.

Other mlk proteins may be tested for mlk ligand binding activity by forming expression of such other proteins as Ig fusion proteins as described herein.

Example 6

Isolation of Full-Length Murine mlk cDNA

The 5' end of the murine mlk can be isolated from random hexamer primed cDNA libraries. Murine skeletal muscle or whole embryo (day 15 post-fertilization) RNA is prepared and used to synthesize double-stranded cDNA. Priming of first strand cDNA with random hexamers rather than oligo dT_{12-18} will ensure the representation of the 5' end of mlk in the library. The cDNA can be ligated with *Eco* RI adaptors and cloned into an EcoRI digested bacteriophage λ vector such as λ ZIPLOX (Gibco-BRL, MD) or λ ZAPII (Stratagene, CA). The resulting plaques can be screened with a radiolabelled synthetic oligonucleotide probe, derived from the 5' end of the available partial cDNA sequence. Alternatively, the library can be screened with a 300-400 bp restriction fragment from the 5' end of the available cDNA.

A murine Balb/c strain skeletal muscle cDNA library (oligo dT and random

hexamer primed) in λgt 10 was used (purchased from Clontech, CA; catalog number ML 3006a). A 5' fragment of pred2#2 was isolated using PCR with primers:

5' ACACTGCGTGGAATGAGCTGA 3' (SEQ ID NO:16) and

5' AAATATGGCAGTCTTGTGCA 3' (SED ID NO:17). This isolated PCR fragment was labelled with α-^{32}P-dCTP and the Klenow fragment of DNA polymerase I with random hexamer priming. The labelled fragment was used to screen the λgt 10 skeletal muscle library by hybridization at high stringency. Single hybridizing bacteriophage were purified. Lambda DNA was purified with Wizard Lambda reagents (Promega, WI), and digested with EcoRI. An EcoRI insert fragment was subcloned into pBluescriptIIKS+ (Stratagene, CA), and examined by nucleotide sequencing. A cDNA clone called "5'mlk #3" was identified by this method. The nucleotide sequence of 5'mlk#3 is provided as SEQ ID NO:14. The amino acid sequence encoded thereby is provided as SEQ ID NO:15.

5'mlk#3 was deposited as an EcoRI insert in pBluescriptIIKS+ with ATCC on January 11, 1995 as accession number 69741.

5'mlk#3 is a partial cDNA for the full length mlk RTK which includes the 5' end. The 5'mlk#3 sequence and previous red2#2 sequence overlap in a region of the extracellular region in juxtaposition to the transmembrane domain. The overlap region runs from about amino acids 331-478 of SEQ ID NO:15 (5'mlk#3) and from about amino acids 1-140 of SEQ ID NO:2 (red2#2).

Alignment of the sequences reveals a 24 bp nucleotide insertion of 5'mlk#3 that is not present in red2#2. This insert encodes an additional 8 amino acids in frame (amino acids 454-461 of SEQ ID NO:15). The first residue after the insert (amino acid 462 of SEQ ID

NO:15) also differs from that of red2#2 (amino acid 124 of SEQ ID NO:2). Thus, there are two isoforms of the mlk RTK, which are designated as "mlk-1" and "mlk-2", which differ by not containing or containing this insert, respectively. Similar isoforms of the KIT and TRK A RTKs have been described in the literature (Crosier et al., Blood, 1993, 82, 1151; Reith et al., EMBO J., 1991, 10, 2451; Barker et al., J. Biol. Chem., 1993, 268, 15150)). Differences in the isoforms may relate to ligand binding, receptor dimerization or activation.

Full-length cDNAs for both isoforms are made by ligating portions of the 5'mlk#3 and red2#2 sequences. Briefly, both cDNAs are cleaved at the same point in the region of the sequence overlap with an appropriate restriction enzyme. The 3' end of the 5'mlk#3 fragment containing the 5' end of the entire mlk coding sequence is then ligated to the 5' end of the red2#2 fragment containing the 3' end of the entire mlk coding sequence.

The nucleotide and amino acid sequences for full-length mlk-1 isoform are provided at SEQ ID NO:18 and SEQ ID NO:19, respectively. The nucleotide and amino acid sequences for full-length mlk-2 isoform are provided at SEQ ID NO:20 and SEQ ID NO:21, respectively. In SEQ ID NO:21 (mlk-2), the 8 amino acid insert is found at amino acids 454-461 and amino acid 462 is a threonine residue (the corresponding residue in SEQ ID NO:19 (mlk-1) is alanine).

Table 1 identified the locations of the mlk protein domains in various forms of mlk protein identified herein. Domains are identified by reference to amino acid numbers in the indicated sequences. The location of the leader sequence aand transmembrane domains was predicted using computer sequence analysis algorithms (leader sequence: Von Heijne, 1987, Sequence Analysis in Molecular Bioilogy: Treasure Trove or Trivial Pursuit (New York:

Academic Press); transmembrane domain: Klein et al., 1985, Biochim. Biophys. Acta, 815, 468).

Table 1
Identification of mlk Protein Domains

	SEQ ID NO:2	SEQ ID NO:15	SEQ ID NO:19	SEQ ID NO:21
leader	N/A	1-21	1-21	1-21
extracellular	1-156	22-478	22-486	22-494
transmembrane	157-177	N/A	487-507	495-515
intracellular	178-530	N/A	508-860	516-868
soluble	1-125; 1-156	N/A	22-453; 22-486	22-462; 22-494
tyrosine kinase	242-517	N/A	572-847	580-855

Example 7

Isolation of Human mlk cDNA

The human mlk cDNA can be isolated by cross-hybridization with the murine mlk cDNA. There is a high degree of cross-species sequence conservation between RTK genes, especially in the region encoding the kinase domain. An example is the similarity between murine mlk and the *Torpedo* RTK cDNA sequences (see Fig. 1). The *Torpedo* RTK nucleotide and amino acid sequences are reported as SEQ ID NO:12 and SEQ ID NO:13, respectively. Human cDNA libraries in appropriate bacteriophage λ vectors such as λ ZAPII can be hybridized with mlk cDNA probes. These libraries can be purchased from Stratagene, California and Clontech, California. The commercially available libraries include human muscle cDNA libraries that would be expected to be a good source. If necessary, the complete human

mlk gene can be isolated from human genomic libraries. Such genomic clones would also provide ideal probes for screening human cDNA libraries at high stringency.

The human cDNA libraries are hybridized with a high specific activity ^{32}P-labelled mlk cDNA probe at low stringency (6 x SSC, 0.1% SDS; 45°C) for 36-48 hours using approximately 2 x 10^6 cpm/ml of probe. The plaque DNA is immobilized on nitrocellulose filters, with approximately 20,000 plaques per 137 mm plate. After the hybridization, the filters are washed in four changes of 4 X SSC, 0.1% SDS at room temperature over 90 minutes, with gentle rotation. Then, the filters are placed in 4 X SSC, 0.1% SDS at 55°C for 30 minutes. The filters are air dried and exposed to x-ray film with two intensifying screens at -80°C, for 24-72 hours. Positive plaques are isolated, re-plated and screened by hybridization until a single, pure recombinant bacteriophage plaque that hybridizes to the probe is isolated, and then analyzed by nucleotide sequencing.

The stringency of the hybridization and filter washing may be increased or decreased to optimize signal detection and signal to background ratio. The exact conditions for cross-hybridization of murine mlk cDNA annealing to the human mlk sequence may also be refined by examining the hybridization of the radiolabelled mlk probe to human genomic DNA Southern blots. Identical Southern blot filters containing *Eco* RI and *Hind* III digested human genomic DNA are hybridized at 5 X SSC, 0.1% SDS at 45°C, and then washed in 5 X SSC, 0.1% SDS at room temperature for 90 minutes. A range of different washing stringencies are then employed over a range of SSC concentrations and temperature conditions. Each wash includes 0.1% SDS, and is carried out for 30 minutes (two changes). In this way, optimal conditions can be found, and then used for cDNA library screening.

An alternative way of isolating the human mlk cDNA is to amplify a portion of it with degenerate primers based on homologies inferred from alignment of murine mlk and *Torpedo* RTK. The template for this PCR amplification is first strand cDNA synthesized from human poly A^+ RNA (purchased from Clontech, CA) bacteriophage λ cDNA libraries or human genomic DNA (Promega, WI). The resulting PCR product is cloned and sequenced, and then represents a probe for hybridizing human cDNA and genomic libraries at high stringency. The unique sequence can also be used to isolate the remainder of the human mlk cDNA by 5' or 3' anchored PCR (Loh et al., Science, 1989, 243, 217; Ohara et al., Proc. Natl. Acad. Sci. USA. 1989, 86, 5673; reviewed in Arnheim et al., Annu. Rev. Biochem., 1992, 61, 131).

Example 8

Expression of Murine and Human mlk cDNA

The mlk-1 cDNA is ligated into the pcDNA3 expression vector (Invitrogen, CA), such that the mlk coding region was under the control of the CMV promoter. This vector also contains an SV40 origin, and can be used as a replicating vector in COS cells. COS cells are transiently transfected with pcDNA3/mlk-1 or as negative controls: an expression vector for Flt4 or no DNA ("mock").

After 48 hours, cell lysates are prepared in ice-cold 1% (v/v) Nonidet P-40, 20 mM Tris-HCl pH 8.0, 150 mM sodiumchloride, 10% (v/v) glycerol, 1 mM sodium vanadate, 1 μg/ml leupeptin, 1 mM Pefabloc SC, 1 μg/ml aprotinin, and then exposed to reducing SDS-PAGE, and transferred to a nitrocellulose filter by western blotting. The filter is blocked with 5% (w/v) non-fat dried milk, 0.1% Tween 20 in TBS (20mM Tris-HCl pH 7.5, 150 mM sodium

chloride) overnight at 4°C. The filter is washed in TBS/0.3% Tween 20, and incubated for 1 hr. at ambient temperature with rabbit antisera against the C-terminal mlk peptide. Then the filter is washed in TBS/0.3% Tween 20 and TBS/0.1% Tween 20. Binding of the antisera is detected with goat anti-rabbit IgG conjugated to horseradish peroxidase. After additional washing, the localized peroxidase is detected with Amersham enchanced chemiluminescence reagents and x-ray film.

The mlk protein may also be expressed as follows:

The mlk cDNA is ligated into a number of mammalian cell expression vectors. For transient expression in COS cells, the mlk cDNA is cloned into pED or related plasmids that place the cDNA under the control of the adenovirus major late promoter with SV40 enhancer and origin, and adenovirus VA RNA genes. The same vector is used for gene amplification in Chinese hamster ovary cells (*dhfr-*), using the linked *DHFR* gene to co-amplify mlk expression with selection in reiteratively increasing concentrations of methotrexate (Kaufman, Methods in Enzymol., 1990, 185, 537).

Expression in cell lines such as factor-dependent BaF3, 32D and FDCP-1, and in PC12 and C2C12 is obtained with cytomegalovirus immediate early, Rous sarcoma virus long terminal repeat, or β-actin promotor sequences. Examples of such vectors are pcDNA3 (Invitrogen, CA) and pDR2 (Clontech, CA). Stable transfectants are selected with G418 or hygromycin B, using neo^R or *hyg B* marker genes incorporated into the expression vectors.

Expression of the RTK is monitored by Northern Blot analysis or ribonuclease protection assay of cellular RNA from transfectants. Antibodies to the murine or human protein can also be prepared and used to examine expression, by immunoprecipitation of biosynthetically

radiolabelled cell extracts, Western blot analysis of cell extracts, or flow cytometry.

Example 9

Identification of mlk Binding Ligands

A number of techniques can be employed to isolate mlk ligand cDNA. Whether the ligand protein is secreted from the cell in a soluble form, or associated with the cell membrane by means of a transmembrane domain, GPI-linked anchor or heparin binding, will determine which route is most successful.

Soluble ligands are detected in the conditioned media of cell lines. The conditioned media is concentrated by Amicon centriprep units 10-100x, and then used to stimulate proliferation of the BaF3 or other factor-dependent cell lines stably transfected with mlk receptor protein expression vectors. Proliferation can be measured by the incorporation of ^{3}H-thymidine. Positive signals in this assay are dissected using neutralizing antibodies to known growth factors, to ensure the response is due to the novel mlk ligand. Other confirmatory data may then be sought. For example, the conditioned media may be used to assay for phosphotyrosine induction on the mlk receptor (see assay above).

In addition to cells transfected with a mlk receptor protein expression vector, other chimeric receptors may be used. For example, a strong proliferative signal is induced in cells such as BaF3 by hematopoietin receptor family members. The extracellular region of the EGF RTK has been fused to the transmembrane and intracellular region of the erythropoietin receptor (EPO-R) (Pacific et al., J. Biol. Chem., 1994, 269, 1571). In 32D cells, this fusion receptor gives a strong proliferation signal to EGF, that is better than that in cells transfected with vectors

encoding the complete EGF RTK. A mlk extracellular region/EPO-R transmembrane-intracellular region fusion functions similarly in 32D and BaF3 cells, and is used for screening conditioned media for the mlk ligand.

The ligand cDNA from a cell line is isolated by either expression cloning or protein purification. A cDNA library is made from polyA+ RNA from the cell line. For expression cloning, this cDNA is ligated into an efficient COS cell expression vector, such as pXM or pED, and is transformed into *E. coli* DH10B cells (Gibco-BRL) by electroporation. Pools of plasmid transformants (approximately 100-2000) are plated individually on 100 mM L-amp plates. Copies are made on nitrocellulose filters. One copy is soaked in glycerol, and is used to cryopreserve the transformants at -80°C. An additional copy is used to make a plasmid DNA preparation from the pool. Such plasmids are transfected transiently into COS cells using DEAE-dextran, and after 48-72 hr. the COS conditioned media are screened on the bioassay. Once a positive pool is identified, the frozen copy is used to isolate the single positive plasmid by a process of subdivision and grid formation.

The ligand can also be purified using the extracellular region of the receptor. The mlk extracellular region is fused in frame to the hinge-C_H2-C_H3 regions of human immunoglobulin (Ig) $\gamma 1$, and is cloned into pEDΔC. Using transient expression in COS cells, or selection and amplification in stably transfected CHO cells, the soluble mlk-Ig fusion is produced. It is purified from conditioned media using protein A-chromatography. The mlk Ig fusion is then coupled to resins and used to affinity purify the mlk ligand. Other purifications steps such as reverse phase or ion exchange chromatography may also be used.

The purified ligand protein is analyzed for amino acid sequence, and the resulting

data is used to design degenerate oligonucleotide primers. These primers are used to screen a cDNA library from the cell line source, either by hybridization in the presence of tetramethyl ammonium chloride, or by PCR.

Membrane-associated mlk ligands are isolated by a different procedure. The mlk-Ig fusion is used to screen for cell lines expressing the mlk ligand. The mlk-Ig fusion is either per-complexed with protein A-FITC, or bound to the cell surface and then detected in a second staining with goat $F_{(ab')2}$ anti-human Ig-FITC (Southern Biotechnology Associates. Alabama). Binding is detected by flow cytometry. A cDNA library from the cell line source is prepared in the pXM or pEDΔC vectors for expression in COS cells. The ligand cDNA can then be isolated by COS transfection, staining with mlk-Ig:protein A-FITC pre-complexed material, and sorting the brightest population of transfected COS cells, e.g., with a Becton-Dickinson FACS Star. Plasmid DNA is rescued from this brightest population of COS cells by the Hirt procedure (Hirt, J. Mol. Biol., 1967, 26, 365), or by an alkaline-SDS mini-plasmid preparation combined with a phenol/chloroform extraction. Plasmid DNA is transformed into *E. coli* DH10B, a large scale plasmid isolation is then carried out and re-transfected into COS cells. By re-iteratively following this selection procedure, the mlk ligand cDNA is enriched (Yamasahi et al., Science, 1988, 241, 825; Rice et al., Cytometry, 1991, 12, 221). After 3-5 rounds of selection, individual plasmids are analyzed by transfection and mlk-Ig binding. Alternatively, the transfected COS cells are enriched for the mlk ligand by "panning" (Seed et al., Proc. Natl. Acad. Sci. USA, 1987, 84, 3365; Aruffo et al., Proc. Natl. Acad. Sci. USA, 1987. 84, 8573), or by magnetic particle separation (Padmanabhan et al. Analytical Biochem., 1988. 170. 341).

Expression libraries are also transfected into COS cells. The subpopulation of

cells expressing the membrane-bound mlk ligand is identified by staining with mlk-Ig and rabbit anti-human Ig conjugated to alkaline phsophatase and visualized with an insoluble alkaline phosphatase substrate, such as NBT/BCIP (Promega, WI). The positive cell population is scraped from the plate, plasmid is recovered and re-transfected in E. coli, and then transfected into COS cells for rescreening. In this way, single cDNAs expressing the mlk ligand can be isolated. This *in situ* staining technique can also be used to identify pools of clones containing the ligand cDNA, and then the specific clone isolated by reiterative pertition and screening of the pool. By permeabilization of the cell membrane, the mlk-Ig can be used to stain for intracellular mlk ligand in the secretory pathway.

All patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Wood, Clive
Caruso, Anthony

(ii) TITLE OF INVENTION: Novel mlk Receptor Tyrosine Kinases

(iii) NUMBER OF SEQUENCES: 21

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LEGAL AFFAIRS
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Brown, Scott A
(B) REGISTRATION NUMBER: 32,724
(C) REFERENCE/DOCKET NUMBER: GI5234A

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8224
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2208 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 13..1602

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GAATTCAAGC CG AAA GAT GCT CTT GTC TTC TTC AAC ACC TCC TAC CGG          48
              Lys Asp Ala Leu Val Phe Phe Asn Thr Ser Tyr Arg
                1               5                  10

GAC CCC GAG GAC GCC CAG GAG CTG CTG ATC CAC ACT GCG TGG AAT GAG        96
Asp Pro Glu Asp Ala Gln Glu Leu Leu Ile His Thr Ala Trp Asn Glu
          15                  20                  25

CTG AAG GCT GTG AGT CCA CTG TGC CGG CCA GCT GCT GAG GCT CTG CTG       144
Leu Lys Ala Val Ser Pro Leu Cys Arg Pro Ala Ala Glu Ala Leu Leu
      30                  35                  40

TGT AAC CAC CTC TTC CAA GAG TGC AGC CCT GGA GTG GTA CCT ACT CCC       192
Cys Asn His Leu Phe Gln Glu Cys Ser Pro Gly Val Val Pro Thr Pro
 45                  50                  55                  60

ATG CCC ATT TGC AGA GAG TAC TGC CTG GCG GTA AAG GAG CTC TTC TGT       240
Met Pro Ile Cys Arg Glu Tyr Cys Leu Ala Val Lys Glu Leu Phe Cys
                  65                  70                  75

GCA AAG GAA TGG CAG GCA ATG GAA GGA AAG GCC CAC CGG GGC CTC TAC       288
Ala Lys Glu Trp Gln Ala Met Glu Gly Lys Ala His Arg Gly Leu Tyr
            80                  85                  90

AGA TCT GGG ATG CAT CTC CTT CCG GTA CCA GAG TGC AGC AAG CTT CCC       336
Arg Ser Gly Met His Leu Leu Pro Val Pro Glu Cys Ser Lys Leu Pro
          95                  100                 105

AGC ATG CAC CGG GAC CCC ACA GCC TGC ACA AGA CTG CCA TAT TTA GCA       384
Ser Met His Arg Asp Pro Thr Ala Cys Thr Arg Leu Pro Tyr Leu Ala
      110                 115                 120

TTC CCG TCA ATA ACG TCC TCC AGG CCG AGC GCG GAC ATT CCA AAC CTG       432
Phe Pro Ser Ile Thr Ser Ser Arg Pro Ser Ala Asp Ile Pro Asn Leu
125                 130                 135                 140

CCT GCC TCC ACC TCT TCC TTT GCC GTC TCG CCT GCG TAC TCC ATG ACC       480
Pro Ala Ser Thr Ser Ser Phe Ala Val Ser Pro Ala Tyr Ser Met Thr
                145                 150                 155

GTC ATC ATC TCC ATC GTG TCC AGC TTT GCC CTG TTT GCT CTT CTC ACC       528
Val Ile Ile Ser Ile Val Ser Ser Phe Ala Leu Phe Ala Leu Leu Thr
            160                 165                 170

ATC GCT ACT CTC TAT TGC TGC CGA AGG AGG AAA GAA TGG AAA AAT AAG       576
Ile Ala Thr Leu Tyr Cys Cys Arg Arg Arg Lys Glu Trp Lys Asn Lys
          175                 180                 185

AAA AGA GAG TCG ACC GCG GTG ACC CTC ACC ACG TTG CCT TCC GAG CTC       624
Lys Arg Glu Ser Thr Ala Val Thr Leu Thr Thr Leu Pro Ser Glu Leu
      190                 195                 200

CTG CTG GAT AGG CTC CAT CCC AAC CCC ATG TAC CAG AGG ATG CCA CTC       672
Leu Leu Asp Arg Leu His Pro Asn Pro Met Tyr Gln Arg Met Pro Leu
205                 210                 215                 220

CTT CTG AAT CCT AAG TTG CTC AGC CTG GAG TAT CCG AGG AAT AAC ATT       720
Leu Leu Asn Pro Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile
                225                 230                 235
```

```
GAG TAT GTC CGA GAC ATC GGA GAG GGG GCG TTT GGA AGA GTC TTC CAA        768
Glu Tyr Val Arg Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln
            240                 245                 250

GCA AGG GCC CCT GGC TTG CTG CCT TAT GAA CCT TTC ACT ATG GTG GCC        816
Ala Arg Ala Pro Gly Leu Leu Pro Tyr Glu Pro Phe Thr Met Val Ala
        255                 260                 265

GTG AAG ATG CTT AAG GAA GAG GCC TCT GCA GAC ATG CAA GCG GAC TTT        864
Val Lys Met Leu Lys Glu Glu Ala Ser Ala Asp Met Gln Ala Asp Phe
    270                 275                 280

CAG AGG GAG GCG GCC CTC ATG GCA GAG TTT GAC AAC CCC AAC ATC GTG        912
Gln Arg Glu Ala Ala Leu Met Ala Glu Phe Asp Asn Pro Asn Ile Val
285                 290                 295                 300

AAA CTC TTA GGT GTG TGT GCC GTT GGG AAG CCG ATG TGT CTG CTC TTT        960
Lys Leu Leu Gly Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe
                305                 310                 315

GAA TAT ATG GCC TAT GGT GAC CTC AAT GAG TTC CTC CGA AGT ATG TCC       1008
Glu Tyr Met Ala Tyr Gly Asp Leu Asn Glu Phe Leu Arg Ser Met Ser
            320                 325                 330

CCG CAC ACT GTT TGC AGC CTC AGC CAC AGT GAC CTG TCC ACG AGG GCT       1056
Pro His Thr Val Cys Ser Leu Ser His Ser Asp Leu Ser Thr Arg Ala
        335                 340                 345

CGG GTG TCT AGC CCT GGT CCT CCA CCA CTG TCC TGT GCA GAA CAG CTC       1104
Arg Val Ser Ser Pro Gly Pro Pro Pro Leu Ser Cys Ala Glu Gln Leu
    350                 355                 360

TGC ATT GCC AGG CAG GTG GCA GCT GGC ATG GCC TAC CTT TCA GAG CGC       1152
Cys Ile Ala Arg Gln Val Ala Ala Gly Met Ala Tyr Leu Ser Glu Arg
365                 370                 375                 380

AAG TTT GTC CAC CGG GAC TTA GCT ACC AGG AAC TGC CTG GTT GGG GAG       1200
Lys Phe Val His Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu
                385                 390                 395

ACC ATG GTG GTG AAA ATT GCA GAC TTT GGC CTC TCC AGG AAC ATC TAT       1248
Thr Met Val Val Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr
            400                 405                 410

TCC GCA GAC TAC TAC AAA GCT GAT GGA AAT GAC GCC ATC CCT ATC CGC       1296
Ser Ala Asp Tyr Tyr Lys Ala Asp Gly Asn Asp Ala Ile Pro Ile Arg
        415                 420                 425

TGG ATG CCG CCC GAG TCT ATC TTC TAC AAC CGC TAC ACC ACG GAG TCG       1344
Trp Met Pro Pro Glu Ser Ile Phe Tyr Asn Arg Tyr Thr Thr Glu Ser
    430                 435                 440

GAT GTA TGG GCC TAT GGT GTG GTC CTC TGG GAG ATC TTC TCC TAT GGG       1392
Asp Val Trp Ala Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Tyr Gly
445                 450                 455                 460

CTG CAG CCC TAC TAT GGA ATG GCC CAC GAG GAG GTC ATT TAC TAT GTG       1440
Leu Gln Pro Tyr Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val
                465                 470                 475
```

```
AGA GAT GGC AAC ATC CTC GCC TGC CCT GAG AAC TGC CCC TTG GAA CTG      1488
Arg Asp Gly Asn Ile Leu Ala Cys Pro Glu Asn Cys Pro Leu Glu Leu
            480                 485                 490

TAC AAC CTC ATG CGC CTG TGT TGG AGC AAG CTG CCT GCT GAT AGA CCC      1536
Tyr Asn Leu Met Arg Leu Cys Trp Ser Lys Leu Pro Ala Asp Arg Pro
        495                 500                 505

AGC TTC TGC AGT ATC CAC AGG ATC CTG CAG CGC ATG TGC GAG AGA GCA      1584
Ser Phe Cys Ser Ile His Arg Ile Leu Gln Arg Met Cys Glu Arg Ala
    510                 515                 520

GAG GGA ACG GTG GGT GTC TAAAGTTGAC CATTCTCAAA CAACACCCAG             1632
Glu Gly Thr Val Gly Val
525                 530

GAGGCTCTTT TCAGACTGTG AGCTGGAGGA ACCCTACCGC AGAGGCCGTG TAAGATCAGA    1692

TAGGAGGAGT TTAACTCAGA CATCACGTGC CAGTTGATTG TTTGCCAGGA GAAACAGATG    1752

GTGAATATCC CACAGGGTTA AAGAGTCACA TCGAAATAGG TTGGAGATAC AGGCTAGGAA    1812

AGAGGACAGC AGGTAGCTCC TCTCCCTCAC AGGGACCGTT TCTAATATAT ATTGCATAAT    1872

AAGAACATCC CGGTTACGTT CCTACATAAT CCCTCAGAGC GAGATTGCAG TGCTTAGGTT    1932

GAATCCAAAA CTGGATGGGC AACTTCATTT TTAACAGAAG ACATCCTGCC CATTGCAAAA    1992

GCAATGTGTC TTTGTGTATA TTTAGGTAAA GGACTGAAAA CTAAAGATAG GAATCCCTTC    2052

TTCCACCAGT CAAGACACGT GGCAGGGTCT TGCTGTTGTT TAGTTCTTCC TTGCACAGAA    2112

TATGTAACGT TGTATTTGCA TTCTGGAATT GAGTATCTAT TTTACTGATA GACTTTTGAA    2172

GAATAAAAAG TGGAAAGCCT GCAAAAAAAA AAGCTT                              2208
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 530 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Lys Asp Ala Leu Val Phe Phe Asn Thr Ser Tyr Arg Asp Pro Glu Asp
 1               5                  10                  15

Ala Gln Glu Leu Leu Ile His Thr Ala Trp Asn Glu Leu Lys Ala Val
            20                  25                  30

Ser Pro Leu Cys Arg Pro Ala Ala Glu Ala Leu Leu Cys Asn His Leu
        35                  40                  45

Phe Gln Glu Cys Ser Pro Gly Val Val Pro Thr Pro Met Pro Ile Cys
    50                  55                  60

Arg Glu Tyr Cys Leu Ala Val Lys Glu Leu Phe Cys Ala Lys Glu Trp
65                  70                  75                  80
```

```
Gln Ala Met Glu Gly Lys Ala His Arg Gly Leu Tyr Arg Ser Gly Met
                85                  90                  95

His Leu Leu Pro Val Pro Glu Cys Ser Lys Leu Pro Ser Met His Arg
            100                 105                 110

Asp Pro Thr Ala Cys Thr Arg Leu Pro Tyr Leu Ala Phe Pro Ser Ile
        115                 120                 125

Thr Ser Ser Arg Pro Ser Ala Asp Ile Pro Asn Leu Pro Ala Ser Thr
    130                 135                 140

Ser Ser Phe Ala Val Ser Pro Ala Tyr Ser Met Thr Val Ile Ile Ser
145                 150                 155                 160

Ile Val Ser Ser Phe Ala Leu Phe Ala Leu Leu Thr Ile Ala Thr Leu
                165                 170                 175

Tyr Cys Cys Arg Arg Arg Lys Glu Trp Lys Asn Lys Lys Arg Glu Ser
            180                 185                 190

Thr Ala Val Thr Leu Thr Thr Leu Pro Ser Glu Leu Leu Leu Asp Arg
        195                 200                 205

Leu His Pro Asn Pro Met Tyr Gln Arg Met Pro Leu Leu Leu Asn Pro
    210                 215                 220

Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile Glu Tyr Val Arg
225                 230                 235                 240

Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln Ala Arg Ala Pro
                245                 250                 255

Gly Leu Leu Pro Tyr Glu Pro Phe Thr Met Val Ala Val Lys Met Leu
            260                 265                 270

Lys Glu Glu Ala Ser Ala Asp Met Gln Ala Asp Phe Gln Arg Glu Ala
        275                 280                 285

Ala Leu Met Ala Glu Phe Asp Asn Pro Asn Ile Val Lys Leu Leu Gly
    290                 295                 300

Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe Glu Tyr Met Ala
305                 310                 315                 320

Tyr Gly Asp Leu Asn Glu Phe Leu Arg Ser Met Ser Pro His Thr Val
                325                 330                 335

Cys Ser Leu Ser His Ser Asp Leu Ser Thr Arg Ala Arg Val Ser Ser
            340                 345                 350

Pro Gly Pro Pro Pro Leu Ser Cys Ala Glu Gln Leu Cys Ile Ala Arg
        355                 360                 365

Gln Val Ala Ala Gly Met Ala Tyr Leu Ser Glu Arg Lys Phe Val His
    370                 375                 380

Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Thr Met Val Val
385                 390                 395                 400

Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr Ser Ala Asp Tyr
                405                 410                 415
```

Tyr Lys Ala Asp Gly Asn Asp Ala Ile Pro Ile Arg Trp Met Pro Pro
420 425 430

Glu Ser Ile Phe Tyr Asn Arg Tyr Thr Thr Glu Ser Asp Val Trp Ala
435 440 445

Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Tyr Gly Leu Gln Pro Tyr
450 455 460

Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg Asp Gly Asn
465 470 475 480

Ile Leu Ala Cys Pro Glu Asn Cys Pro Leu Glu Leu Tyr Asn Leu Met
485 490 495

Arg Leu Cys Trp Ser Lys Leu Pro Ala Asp Arg Pro Ser Phe Cys Ser
500 505 510

Ile His Arg Ile Leu Gln Arg Met Cys Glu Arg Ala Glu Gly Thr Val
515 520 525

Gly Val
530

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTTAGGATCC ACMGNGAYYT 20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGGAATTCCR WAGGACCASA CRTC 24

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATGAATTCG AYYTNGCNGC NMGNAA 26

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTTGAATTCT CNARNGCNAR CCA 23

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TACTACAAAG CTGATGGAAA TGACG 25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAGACTGTGA GCTGGAGGAC C 21

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGATTATGTA GGAACGTAAC C 21

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGAAATAGGT TGGAGATACA GG 22

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGTGTCTTGA CTGGTGGAAG AAGG 24

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3398 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 121..2961

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CTTTCGGGAC TGTCAGTGAA TCCAGAGAAG CTAACATCTA TGAACTCTGA TTTGAATACA 60

GGATTCAAGC GTGTACTGGC CTGTTTGGCA GAAAAATATC ATTTCTGATC GACGATCATC 120

ATG AAC TTT ATC CCA GTC GAC ATT CCA CTC TTG ATG ATC TTC CTT GTG 168
Met Asn Phe Ile Pro Val Asp Ile Pro Leu Leu Met Ile Phe Leu Val
1 5 10 15

ACA ACT GGG GGC TCA GCT GAC GGA ATC CTT CCC AAA GCT CCA CAG ATC 216
Thr Thr Gly Gly Ser Ala Asp Gly Ile Leu Pro Lys Ala Pro Gln Ile
20 25 30

ACC AGT CCC TTG GAG ACA GTG GAT GCC TTG GTT GAG GAA GAA GCT TCT 264
Thr Ser Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Glu Ala Ser
35 40 45

TTC ATG TGT GCA GTG GAT TCA TAC CCA GCG GCA GAG ATT ACC TGG ACC 312
Phe Met Cys Ala Val Asp Ser Tyr Pro Ala Ala Glu Ile Thr Trp Thr
50 55 60

CGC AAT AAC ATT CCC ATA AGA CCC TTT GAC ACT CGC TAC AGT ACA AAA 360
Arg Asn Asn Ile Pro Ile Arg Pro Phe Asp Thr Arg Tyr Ser Thr Lys
65 70 75 80

GAA AAT GGC CAG ATA TTA ACC ATC CTC AGC GTT GAA GAC ACA GAC AAT 408
Glu Asn Gly Gln Ile Leu Thr Ile Leu Ser Val Glu Asp Thr Asp Asn
85 90 95

GGG GTG TAC TGC TGC ACC GCC AAC AAC GGC ATG GGG AGC TCT GCT CAA 456
Gly Val Tyr Cys Cys Thr Ala Asn Asn Gly Met Gly Ser Ser Ala Gln
100 105 110

AGC TGT GGT GCC CTC CAG GTC AAA ATG AAG CCA AAG ATC ATT CGG CCA 504
Ser Cys Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Ile Arg Pro
115 120 125

```
CCC ACT GAT GTC AGA GCA CTG CTG GGA TCG AAG GTT GTG TTA CCT TGC      552
Pro Thr Asp Val Arg Ala Leu Leu Gly Ser Lys Val Val Leu Pro Cys
    130                 135                 140

AGT ACC ATG GGG AAT CCA AAA CCA GCC ATT TCA TGG TTC AAA GAT GAA      600
Ser Thr Met Gly Asn Pro Lys Pro Ala Ile Ser Trp Phe Lys Asp Glu
145                 150                 155                 160

ACT GCA CTG AAA AAT GAC CAG CCT CGA ACT TCT GTC CTT GAG TCT GGG      648
Thr Ala Leu Lys Asn Asp Gln Pro Arg Thr Ser Val Leu Glu Ser Gly
                165                 170                 175

AAC TTA AGA ATT CGC AAT GTT CAG CTT GAA GAT GCA GGA AAA TAT CGA      696
Asn Leu Arg Ile Arg Asn Val Gln Leu Glu Asp Ala Gly Lys Tyr Arg
            180                 185                 190

TGT TTG GCA AGA AAC AGC CTG GGC TTC GAG TAT TCC AGA TCT GCG GCT      744
Cys Leu Ala Arg Asn Ser Leu Gly Phe Glu Tyr Ser Arg Ser Ala Ala
        195                 200                 205

CTG GAA GTG CAG GTC TCT GCC AGA ATT GTG AAG GCG CCC ACA TCA CAA      792
Leu Glu Val Gln Val Ser Ala Arg Ile Val Lys Ala Pro Thr Ser Gln
    210                 215                 220

AAT GTC AGC TAT GGT TCC GAA GTG ATC TTG CAG TGC AAA GCC ACC GGG      840
Asn Val Ser Tyr Gly Ser Glu Val Ile Leu Gln Cys Lys Ala Thr Gly
225                 230                 235                 240

TTT CCG ATT CCC ACC ATC AAG TGG TTG GAG AAT GGG AGA GCA GTC CCC      888
Phe Pro Ile Pro Thr Ile Lys Trp Leu Glu Asn Gly Arg Ala Val Pro
                245                 250                 255

AAG GGT TCG ATA CAG AAT CGC ATC AAG GGA GAG GTG ATG GAA TCT AGG      936
Lys Gly Ser Ile Gln Asn Arg Ile Lys Gly Glu Val Met Glu Ser Arg
            260                 265                 270

CTG CGG GTC TAT GTT ACC AGA CCT TCA CTG TTC ACT TGC CTG ACT ACC      984
Leu Arg Val Tyr Val Thr Arg Pro Ser Leu Phe Thr Cys Leu Thr Thr
        275                 280                 285

AAC AAG CAC AAT GAA GGA AGT ACC ACA GCA AAA GCC ACT GCC ACC CTG     1032
Asn Lys His Asn Glu Gly Ser Thr Thr Ala Lys Ala Thr Ala Thr Leu
    290                 295                 300

GAT ATC AAA GAA TGG AGA TTG TAC AAA GGT GAC TTG GGC TAT TGC AGC     1080
Asp Ile Lys Glu Trp Arg Leu Tyr Lys Gly Asp Leu Gly Tyr Cys Ser
305                 310                 315                 320

ACA TAT CGT GGT GAG GTA TGC CAA GGT CTT CTG GGA AAT GGC CAG CTG     1128
Thr Tyr Arg Gly Glu Val Cys Gln Gly Leu Leu Gly Asn Gly Gln Leu
                325                 330                 335

GTT TTC TTC AAC TCT TCT TTT GCC GAT GCA GAG GGG ACA CAA GAG ATG     1176
Val Phe Phe Asn Ser Ser Phe Ala Asp Ala Glu Gly Thr Gln Glu Met
            340                 345                 350

ATG GCC AGG AGC ACA TGG ACG GAG TTG GAT GGC GTC AGC TTG CTG TGC     1224
Met Ala Arg Ser Thr Trp Thr Glu Leu Asp Gly Val Ser Leu Leu Cys
        355                 360                 365
```

```
AAA CCA GCT GCC GAG TCC CTA CTC TGC CAC TTC ATT TTC CAA GAC TGT      1272
Lys Pro Ala Ala Glu Ser Leu Leu Cys His Phe Ile Phe Gln Asp Cys
    370                 375                 380

AAT CCT TTA GGG CTG GGT CCT ACT CCC AAA CTT GTG TGC CGT GAG CAT      1320
Asn Pro Leu Gly Leu Gly Pro Thr Pro Lys Leu Val Cys Arg Glu His
385                 390                 395                 400

TGC TTG GCA GTG AAA GAG CTT TAT TGT TAC AAA GAA TGG ATC ACA ATG      1368
Cys Leu Ala Val Lys Glu Leu Tyr Cys Tyr Lys Glu Trp Ile Thr Met
                405                 410                 415

GAG GAC AAT TCA CGC ATA GGA GTT TAC TCT GCG GGT CTG AGC CTA CCA      1416
Glu Asp Asn Ser Arg Ile Gly Val Tyr Ser Ala Gly Leu Ser Leu Pro
            420                 425                 430

GAC TGT CAG AGG CTT CCC AGT ATA CAC CAT GAC CCA GAA GCA TGC ACC      1464
Asp Cys Gln Arg Leu Pro Ser Ile His His Asp Pro Glu Ala Cys Thr
        435                 440                 445

AGA GTC TCT TTT CTT GAC ATG AAG AAG GGG CTC GTT ACC AGA ATG TGT      1512
Arg Val Ser Phe Leu Asp Met Lys Lys Gly Leu Val Thr Arg Met Cys
    450                 455                 460

TAC AAC AAT AAC GGG AGG TTT TAC CAG GGA TCG GTG AAT GTC ACT GCA      1560
Tyr Asn Asn Asn Gly Arg Phe Tyr Gln Gly Ser Val Asn Val Thr Ala
465                 470                 475                 480

TCA GGC ATT TCC TGT CAG AGA TGG AGT GAG CAG GCT CCT CAT TTC CAC      1608
Ser Gly Ile Ser Cys Gln Arg Trp Ser Glu Gln Ala Pro His Phe His
                485                 490                 495

AGG CGT CTG CCA GAG ATA TTT CCT GAA TTA GCC AAT TCT GAC AAC TTC      1656
Arg Arg Leu Pro Glu Ile Phe Pro Glu Leu Ala Asn Ser Asp Asn Phe
            500                 505                 510

TGC CGG AAC CCA GGA GGT GAG AGT GAA CGA CCG TGG TGT TAT ACG ATG      1704
Cys Arg Asn Pro Gly Gly Glu Ser Glu Arg Pro Trp Cys Tyr Thr Met
        515                 520                 525

GAT CGA GAC ATC CGG TGG GAA TTC TGC AAT GTG CCT CAA TGT ATC AAT      1752
Asp Arg Asp Ile Arg Trp Glu Phe Cys Asn Val Pro Gln Cys Ile Asn
    530                 535                 540

GTT TCC TCA ATA TCA GAG ATG AAG CCT AAA ACA GAA ACA GCC AAC ACT      1800
Val Ser Ser Ile Ser Glu Met Lys Pro Lys Thr Glu Thr Ala Asn Thr
545                 550                 555                 560

CCC AGC ACT TCT GCC ACC TAC TCA ATG ACC GTC ATA ATT TCC ATA ATC      1848
Pro Ser Thr Ser Ala Thr Tyr Ser Met Thr Val Ile Ile Ser Ile Ile
                565                 570                 575

TCC AGC CTT GCA GCC TCC ATC CTG TTG ATA ATT ATA ATT CTC ACT TGT      1896
Ser Ser Leu Ala Ala Ser Ile Leu Leu Ile Ile Ile Ile Leu Thr Cys
            580                 585                 590

CAC CAT CAC CAG AAG GGA TTG CAG ACC AGA AAG AGT TAC AGA ACA ACT      1944
His His His Gln Lys Gly Leu Gln Thr Arg Lys Ser Tyr Arg Thr Thr
        595                 600                 605
```

```
GAG ACC CCT ACT CTG GCT ACT CTT CCT TCA GAG CTG CTT CTA GAC AGA      1992
Glu Thr Pro Thr Leu Ala Thr Leu Pro Ser Glu Leu Leu Leu Asp Arg
    610                 615                 620

CTT CAC CCC AAC CCA ATG TAC CAG CGC CTG CCT CTT CTT CTA AAT GCT      2040
Leu His Pro Asn Pro Met Tyr Gln Arg Leu Pro Leu Leu Leu Asn Ala
625                 630                 635                 640

AAA CTA CTG AGC CTC GAG TAT CCA AGG AAT AAC ATA GAA TAT GTG CGG      2088
Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile Glu Tyr Val Arg
                645                 650                 655

GAT ATT GGA GAG GGA GCA TTT GGA AGA GTA TTC CAG GCA AGA GCC CCT      2136
Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln Ala Arg Ala Pro
            660                 665                 670

CAT CTG CTG CCG CAG GAG ACC TCC ACC ATG GTG GCT GTG AAG ATG CTT      2184
His Leu Leu Pro Gln Glu Thr Ser Thr Met Val Ala Val Lys Met Leu
        675                 680                 685

AAA GAA GAA GCG TCA CCT GAC ATG CAG GCA GAC TTC CGG AGA GAA GCA      2232
Lys Glu Glu Ala Ser Pro Asp Met Gln Ala Asp Phe Arg Arg Glu Ala
    690                 695                 700

GCG CTC ATG GCA GAG TTC AAC CAT CCA AAC ATC GTC AAG CTT TTA GGA      2280
Ala Leu Met Ala Glu Phe Asn His Pro Asn Ile Val Lys Leu Leu Gly
705                 710                 715                 720

GTG TGC GCT GTT GGA AAG CCG ATG TGC CTG CTA TTC GAG TAC ATG GCG      2328
Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe Glu Tyr Met Ala
                725                 730                 735

CAT GGA GAC CTG AAC GAG TAT TTA CGC AAG CGG TCA CCC ATC ACC GCC      2376
His Gly Asp Leu Asn Glu Tyr Leu Arg Lys Arg Ser Pro Ile Thr Ala
            740                 745                 750

CGC ACC TTG AGG CCC GCC AAT TGT GTG GGA TGG AGC AGC GGC TGG GGA      2424
Arg Thr Leu Arg Pro Ala Asn Cys Val Gly Trp Ser Ser Gly Trp Gly
        755                 760                 765

AAG GGC CTG ACA GCC CTC AGC TGC GCT GAC CAA CTG AAC ATC GCC AAG      2472
Lys Gly Leu Thr Ala Leu Ser Cys Ala Asp Gln Leu Asn Ile Ala Lys
    770                 775                 780

CAG ATC TCA GCG GGC ATG ACC TAC CTG TCG GAG CGC AAG TTT GTT CAC      2520
Gln Ile Ser Ala Gly Met Thr Tyr Leu Ser Glu Arg Lys Phe Val His
785                 790                 795                 800

CGG GAC CTG GCC ACC CGT AAC TGC TTG GTT GGA GAG AAG CTG GTA GTT      2568
Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Lys Leu Val Val
                805                 810                 815

AAG ATT GCT GAC TTT GGC CTC TCC AGG AAC ATC TAC TCT GCG GAC TAT      2616
Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr Ser Ala Asp Tyr
            820                 825                 830

TAC AAG GCC AAT GAG AAT GAT GCC ATC CCG ATC AGG TGG ATG CCT CCT      2664
Tyr Lys Ala Asn Glu Asn Asp Ala Ile Pro Ile Arg Trp Met Pro Pro
        835                 840                 845
```

```
GAA TCC ATA TTC TTC AAC CGT TAT ACC ACC GAG TCC GAC GTC TGG GCT      2712
Glu Ser Ile Phe Phe Asn Arg Tyr Thr Thr Glu Ser Asp Val Trp Ala
    850                 855                 860

TAT GGT GTG GTC CTG TGG GAG ATC TTC TCG TCC GGC ATG CAG CCA TAC      2760
Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Ser Gly Met Gln Pro Tyr
865                 870                 875                 880

TAT GGG ATG GCC CAC GAA GAG GTG ATC TAC TAT GTT CGA GAC GGG AAC      2808
Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg Asp Gly Asn
                885                 890                 895

ATC CTG TCC TGC CCC GAG AAC TGC CCA CCA GAG CTG TAC AAC TTG ATG      2856
Ile Leu Ser Cys Pro Glu Asn Cys Pro Pro Glu Leu Tyr Asn Leu Met
            900                 905                 910

CGC CTC TGC TGG AGT AAC ATG CCA TCA GAC AGA CCG ACG TTC GCC AGT      2904
Arg Leu Cys Trp Ser Asn Met Pro Ser Asp Arg Pro Thr Phe Ala Ser
        915                 920                 925

ATC CAT CGC ATC CTG GAG CGC ATG CAC CAG AGG ATG GCA GCC GCA CTC      2952
Ile His Arg Ile Leu Glu Arg Met His Gln Arg Met Ala Ala Ala Leu
    930                 935                 940

CCA GTC TGATCCCCCC TCCCCCCTCT GTACCTTGGG GTACATGTTC CCTGTGCAAG       3008
Pro Val
945

ATTGGCAAGG GTTATGGCCG GTTGCCTGAC CGTGCCTGAC TTCGGATGCC AGGAGCACAC    3068

ACCACGCCAA CCAGGCCCTT AAATACTCTG ACCCCCACAA TCTTCATTTC ACAAACTAGC    3128

CATGCCCAGT ACATTGAAAC ACCCAATATT AAACCGTAGG TTCCAATTAC CATCCCTCCC    3188

ATTTTTTGTT CATGAACACG TTTCTCAATT ATTACTGGAT GTTCATTTAA TTATATATGA    3248

TGACATTTTA TGAGGGGCTT CAGTAAAGAA GACAGGGAGA CTGTCACCAT GGTAGAAGGC    3308

TCGTAACCAA AGGCCCCTAA TTTAAAAAGA AAAGTCAGGG GATTTTTTTT TTTGTGGATT    3368

GTTCCAATCA ACTTGAAAAA GAAGGAATTC                                     3398
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 946 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Asn Phe Ile Pro Val Asp Ile Pro Leu Leu Met Ile Phe Leu Val
  1               5                  10                  15

Thr Thr Gly Gly Ser Ala Asp Gly Ile Leu Pro Lys Ala Pro Gln Ile
             20                  25                  30

Thr Ser Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Glu Ala Ser
         35                  40                  45
```

```
Phe Met Cys Ala Val Asp Ser Tyr Pro Ala Ala Glu Ile Thr Trp Thr
     50                  55                  60

Arg Asn Asn Ile Pro Ile Arg Pro Phe Asp Thr Arg Tyr Ser Thr Lys
 65                  70                  75                  80

Glu Asn Gly Gln Ile Leu Thr Ile Leu Ser Val Glu Asp Thr Asp Asn
                 85                  90                  95

Gly Val Tyr Cys Cys Thr Ala Asn Asn Gly Met Gly Ser Ser Ala Gln
            100                 105                 110

Ser Cys Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Ile Arg Pro
        115                 120                 125

Pro Thr Asp Val Arg Ala Leu Leu Gly Ser Lys Val Val Leu Pro Cys
    130                 135                 140

Ser Thr Met Gly Asn Pro Lys Pro Ala Ile Ser Trp Phe Lys Asp Glu
145                 150                 155                 160

Thr Ala Leu Lys Asn Asp Gln Pro Arg Thr Ser Val Leu Glu Ser Gly
                165                 170                 175

Asn Leu Arg Ile Arg Asn Val Gln Leu Glu Asp Ala Gly Lys Tyr Arg
            180                 185                 190

Cys Leu Ala Arg Asn Ser Leu Gly Phe Glu Tyr Ser Arg Ser Ala Ala
        195                 200                 205

Leu Glu Val Gln Val Ser Ala Arg Ile Val Lys Ala Pro Thr Ser Gln
    210                 215                 220

Asn Val Ser Tyr Gly Ser Glu Val Ile Leu Gln Cys Lys Ala Thr Gly
225                 230                 235                 240

Phe Pro Ile Pro Thr Ile Lys Trp Leu Glu Asn Gly Arg Ala Val Pro
                245                 250                 255

Lys Gly Ser Ile Gln Asn Arg Ile Lys Gly Glu Val Met Glu Ser Arg
            260                 265                 270

Leu Arg Val Tyr Val Thr Arg Pro Ser Leu Phe Thr Cys Leu Thr Thr
        275                 280                 285

Asn Lys His Asn Glu Gly Ser Thr Thr Ala Lys Ala Thr Ala Thr Leu
    290                 295                 300

Asp Ile Lys Glu Trp Arg Leu Tyr Lys Gly Asp Leu Gly Tyr Cys Ser
305                 310                 315                 320

Thr Tyr Arg Gly Glu Val Cys Gln Gly Leu Leu Gly Asn Gly Gln Leu
                325                 330                 335

Val Phe Phe Asn Ser Ser Phe Ala Asp Ala Glu Gly Thr Gln Glu Met
            340                 345                 350

Met Ala Arg Ser Thr Trp Thr Glu Leu Asp Gly Val Ser Leu Leu Cys
        355                 360                 365

Lys Pro Ala Ala Glu Ser Leu Leu Cys His Phe Ile Phe Gln Asp Cys
    370                 375                 380
```

```
Asn Pro Leu Gly Leu Gly Pro Thr Pro Lys Leu Val Cys Arg Glu His
385                 390                 395                 400

Cys Leu Ala Val Lys Glu Leu Tyr Cys Tyr Lys Glu Trp Ile Thr Met
                405                 410                 415

Glu Asp Asn Ser Arg Ile Gly Val Tyr Ser Ala Gly Leu Ser Leu Pro
            420                 425                 430

Asp Cys Gln Arg Leu Pro Ser Ile His His Asp Pro Glu Ala Cys Thr
        435                 440                 445

Arg Val Ser Phe Leu Asp Met Lys Lys Gly Leu Val Thr Arg Met Cys
    450                 455                 460

Tyr Asn Asn Asn Gly Arg Phe Tyr Gln Gly Ser Val Asn Val Thr Ala
465                 470                 475                 480

Ser Gly Ile Ser Cys Gln Arg Trp Ser Glu Gln Ala Pro His Phe His
                485                 490                 495

Arg Arg Leu Pro Glu Ile Phe Pro Glu Leu Ala Asn Ser Asp Asn Phe
            500                 505                 510

Cys Arg Asn Pro Gly Gly Glu Ser Glu Arg Pro Trp Cys Tyr Thr Met
        515                 520                 525

Asp Arg Asp Ile Arg Trp Glu Phe Cys Asn Val Pro Gln Cys Ile Asn
    530                 535                 540

Val Ser Ser Ile Ser Glu Met Lys Pro Lys Thr Glu Thr Ala Asn Thr
545                 550                 555                 560

Pro Ser Thr Ser Ala Thr Tyr Ser Met Thr Val Ile Ile Ser Ile Ile
                565                 570                 575

Ser Ser Leu Ala Ala Ser Ile Leu Leu Ile Ile Ile Ile Leu Thr Cys
            580                 585                 590

His His His Gln Lys Gly Leu Gln Thr Arg Lys Ser Tyr Arg Thr Thr
        595                 600                 605

Glu Thr Pro Thr Leu Ala Thr Leu Pro Ser Glu Leu Leu Leu Asp Arg
    610                 615                 620

Leu His Pro Asn Pro Met Tyr Gln Arg Leu Pro Leu Leu Leu Asn Ala
625                 630                 635                 640

Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile Glu Tyr Val Arg
                645                 650                 655

Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln Ala Arg Ala Pro
            660                 665                 670

His Leu Leu Pro Gln Glu Thr Ser Thr Met Val Ala Val Lys Met Leu
        675                 680                 685

Lys Glu Glu Ala Ser Pro Asp Met Gln Ala Asp Phe Arg Arg Glu Ala
    690                 695                 700

Ala Leu Met Ala Glu Phe Asn His Pro Asn Ile Val Lys Leu Leu Gly
705                 710                 715                 720
```

```
Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe Glu Tyr Met Ala
                725                 730                 735

His Gly Asp Leu Asn Glu Tyr Leu Arg Lys Arg Ser Pro Ile Thr Ala
            740                 745                 750

Arg Thr Leu Arg Pro Ala Asn Cys Val Gly Trp Ser Ser Gly Trp Gly
        755                 760                 765

Lys Gly Leu Thr Ala Leu Ser Cys Ala Asp Gln Leu Asn Ile Ala Lys
    770                 775                 780

Gln Ile Ser Ala Gly Met Thr Tyr Leu Ser Glu Arg Lys Phe Val His
785                 790                 795                 800

Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Lys Leu Val Val
                805                 810                 815

Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr Ser Ala Asp Tyr
            820                 825                 830

Tyr Lys Ala Asn Glu Asn Asp Ala Ile Pro Ile Arg Trp Met Pro Pro
        835                 840                 845

Glu Ser Ile Phe Phe Asn Arg Tyr Thr Thr Glu Ser Asp Val Trp Ala
    850                 855                 860

Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Ser Gly Met Gln Pro Tyr
865                 870                 875                 880

Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg Asp Gly Asn
                885                 890                 895

Ile Leu Ser Cys Pro Glu Asn Cys Pro Pro Glu Leu Tyr Asn Leu Met
            900                 905                 910

Arg Leu Cys Trp Ser Asn Met Pro Ser Asp Arg Pro Thr Phe Ala Ser
        915                 920                 925

Ile His Arg Ile Leu Glu Arg Met His Gln Arg Met Ala Ala Ala Leu
    930                 935                 940

Pro Val
945
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1581 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 146..1579

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
CAGACACAAA CAGTCCTTAG CGGACAACTC TATTGTAACA AACCATGCTT TAAAATGTAA      60

ACCCGGGAGC GTGTTTTTTG TTTTTTTTTT TTTTTTTTCC TCACGTTGTC CAGAAGCAAC     120

CTTTCTTCCT GAGCCTGGAT TAATC ATG AGA GAG CTT GTC AAC ATT CCA CTG        172
                            Met Arg Glu Leu Val Asn Ile Pro Leu
                              1               5

TTA CAG ATG CTC ACC CTG GTT GCC TTC AGC GGG ACT GAG AAA CTT CCA        220
Leu Gln Met Leu Thr Leu Val Ala Phe Ser Gly Thr Glu Lys Leu Pro
 10                  15                  20                  25

AAA GCC CCT GTC ATC ACC ACG CCT CTT GAA ACT GTA GAT GCC TTG GTT        268
Lys Ala Pro Val Ile Thr Thr Pro Leu Glu Thr Val Asp Ala Leu Val
                30                  35                  40

GAA GAA GTA GCG ACT TTC ATG TGT GCC GTG GAA TCC TAC CCT CAG CCC        316
Glu Glu Val Ala Thr Phe Met Cys Ala Val Glu Ser Tyr Pro Gln Pro
            45                  50                  55

GAG ATT TCT TGG ACC AGA AAT AAA ATT CTC ATT AAG CTG TTT GAC ACC        364
Glu Ile Ser Trp Thr Arg Asn Lys Ile Leu Ile Lys Leu Phe Asp Thr
        60                  65                  70

CGC TAC AGC ATC CGG GAG AAT GGT CAG CTC CTC ACC ATT CTG AGC GTG        412
Arg Tyr Ser Ile Arg Glu Asn Gly Gln Leu Leu Thr Ile Leu Ser Val
    75                  80                  85

GAA GAC AGT GAT GAT GGC ATC TAC TGC TGC ATA GCC AAC AAT GGA GTG        460
Glu Asp Ser Asp Asp Gly Ile Tyr Cys Cys Ile Ala Asn Asn Gly Val
 90                  95                  100                 105

GGA GGA GCC GTG GAG AGT TGT GGT GCC CTG CAA GTG AAG ATG AAA CCT        508
Gly Gly Ala Val Glu Ser Cys Gly Ala Leu Gln Val Lys Met Lys Pro
                110                 115                 120

AAA ATA ACT CGT CCT CCC ATT AAT GTA AAA ATA ATA GAG GGA TTG AAG        556
Lys Ile Thr Arg Pro Pro Ile Asn Val Lys Ile Ile Glu Gly Leu Lys
            125                 130                 135

GCA GTT CTG CCG TGC ACT ACG ATG GGT AAC CCC AAA CCA TCT GTG TCC        604
Ala Val Leu Pro Cys Thr Thr Met Gly Asn Pro Lys Pro Ser Val Ser
        140                 145                 150

TGG ATC AAG GGG GAC AAT GCT CTC AGG GAA AAT TCC AGA ATC GCA GTT        652
Trp Ile Lys Gly Asp Asn Ala Leu Arg Glu Asn Ser Arg Ile Ala Val
    155                 160                 165

CTT GAA TCT GGG AGC TTA AGG ATC CAT AAT GTG CAA AAG GAA GAT GCA        700
Leu Glu Ser Gly Ser Leu Arg Ile His Asn Val Gln Lys Glu Asp Ala
170                 175                 180                 185

GGA CAG TAC CGC TGT GTG GCC AAA AAC AGC CTG GGC ACA GCT TAC TCC        748
Gly Gln Tyr Arg Cys Val Ala Lys Asn Ser Leu Gly Thr Ala Tyr Ser
                190                 195                 200

AAA CTG GTG AAG CTG GAA GTG GAG GTT TTT GCA AGA ATC CTG CGT GCT        796
Lys Leu Val Lys Leu Glu Val Glu Val Phe Ala Arg Ile Leu Arg Ala
            205                 210                 215
```

```
CCT GAA TCC CAC AAT GTC ACC TTT GGT TCC TTT GTA ACC CTA CGC TGC        844
Pro Glu Ser His Asn Val Thr Phe Gly Ser Phe Val Thr Leu Arg Cys
        220                 225                 230

ACA GCA ATA GGC ATC CCT GTC CCC ACC ATC AGC TGG ATT GAA AAC GGA        892
Thr Ala Ile Gly Ile Pro Val Pro Thr Ile Ser Trp Ile Glu Asn Gly
    235                 240                 245

AAT GCT GTT TCT TCA GGT TCC ATT CAA GAG AGT GTG AAA GAC CGA GTG        940
Asn Ala Val Ser Ser Gly Ser Ile Gln Glu Ser Val Lys Asp Arg Val
250                 255                 260                 265

ATT GAC TCA AGA CTC CAG CTC TTC ATC ACA AAG CCA GGA CTC TAC ACA        988
Ile Asp Ser Arg Leu Gln Leu Phe Ile Thr Lys Pro Gly Leu Tyr Thr
                270                 275                 280

TGC ATA GCT ACC AAT AAG CAC GGA GAA AAG TTC AGT ACC GCA AAG GCT       1036
Cys Ile Ala Thr Asn Lys His Gly Glu Lys Phe Ser Thr Ala Lys Ala
            285                 290                 295

GCA GCC ACT GTC AGC ATA GCA GAA TGG AGT AAG TCA CAG AAA GAC AGC       1084
Ala Ala Thr Val Ser Ile Ala Glu Trp Ser Lys Ser Gln Lys Asp Ser
        300                 305                 310

CAA GGC TAC TGT GCC CAG TAC AGA GGG GAG GTG TGT GAT GCA GTC CTG       1132
Gln Gly Tyr Cys Ala Gln Tyr Arg Gly Glu Val Cys Asp Ala Val Leu
    315                 320                 325

GCG AAA GAT GCT CTT GTC TTC TTC AAC ACC TCC TAC CGG GAC CCC GAG       1180
Ala Lys Asp Ala Leu Val Phe Phe Asn Thr Ser Tyr Arg Asp Pro Glu
330                 335                 340                 345

GAC GCC CAG GAG CTG CTG ATC CAC ACT GCG TGG AAT GAG CTG AAG GCT       1228
Asp Ala Gln Glu Leu Leu Ile His Thr Ala Trp Asn Glu Leu Lys Ala
                350                 355                 360

GTG AGT CCA CTG TGC CGG CCA GCT GCT GAG GCT CTG CTG TGT AAC CAC       1276
Val Ser Pro Leu Cys Arg Pro Ala Ala Glu Ala Leu Leu Cys Asn His
            365                 370                 375

CTC TTC CAA GAG TGC AGC CCT GGA GTG GTA CCT ACT CCC ATG CCC ATT       1324
Leu Phe Gln Glu Cys Ser Pro Gly Val Val Pro Thr Pro Met Pro Ile
        380                 385                 390

TGC AGA GAG TAC TGC CTG GCG GTA AAG GAG CTC TTC TGT GCA AAG GAA       1372
Cys Arg Glu Tyr Cys Leu Ala Val Lys Glu Leu Phe Cys Ala Lys Glu
    395                 400                 405

TGG CAG GCA ATG GAA GGA AAG GCC CAC CGG GGC CTC TAC AGA TCT GGG       1420
Trp Gln Ala Met Glu Gly Lys Ala His Arg Gly Leu Tyr Arg Ser Gly
410                 415                 420                 425

ATG CAT CTC CTT CCG GTA CCA GAG TGC AGC AAG CTT CCC AGC ATG CAC       1468
Met His Leu Leu Pro Val Pro Glu Cys Ser Lys Leu Pro Ser Met His
                430                 435                 440

CGG GAC CCC ACA GCC TGC ACA AGA CTG CCA TAT TTA GAT TAT AAA AAA       1516
Arg Asp Pro Thr Ala Cys Thr Arg Leu Pro Tyr Leu Asp Tyr Lys Lys
            445                 450                 455
```

```
GAA AAC ATA ACA ACA TTC CCG TCA ATA ACG TCC TCC AGG CCG AGC GCG      1564
Glu Asn Ile Thr Thr Phe Pro Ser Ile Thr Ser Ser Arg Pro Ser Ala
        460                 465                 470

GAC ATT CCA AAC CTG CC                                           1581
Asp Ile Pro Asn Leu
    475
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 478 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Arg Glu Leu Val Asn Ile Pro Leu Leu Gln Met Leu Thr Leu Val
 1               5                  10                  15

Ala Phe Ser Gly Thr Glu Lys Leu Pro Lys Ala Pro Val Ile Thr Thr
            20                  25                  30

Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Val Ala Thr Phe Met
        35                  40                  45

Cys Ala Val Glu Ser Tyr Pro Gln Pro Glu Ile Ser Trp Thr Arg Asn
    50                  55                  60

Lys Ile Leu Ile Lys Leu Phe Asp Thr Arg Tyr Ser Ile Arg Glu Asn
65                  70                  75                  80

Gly Gln Leu Leu Thr Ile Leu Ser Val Glu Asp Ser Asp Asp Gly Ile
                85                  90                  95

Tyr Cys Cys Ile Ala Asn Asn Gly Val Gly Gly Ala Val Glu Ser Cys
            100                 105                 110

Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Thr Arg Pro Pro Ile
        115                 120                 125

Asn Val Lys Ile Ile Glu Gly Leu Lys Ala Val Leu Pro Cys Thr Thr
    130                 135                 140

Met Gly Asn Pro Lys Pro Ser Val Ser Trp Ile Lys Gly Asp Asn Ala
145                 150                 155                 160

Leu Arg Glu Asn Ser Arg Ile Ala Val Leu Glu Ser Gly Ser Leu Arg
                165                 170                 175

Ile His Asn Val Gln Lys Glu Asp Ala Gly Gln Tyr Arg Cys Val Ala
            180                 185                 190

Lys Asn Ser Leu Gly Thr Ala Tyr Ser Lys Leu Val Lys Leu Glu Val
        195                 200                 205

Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr
    210                 215                 220
```

```
Phe Gly Ser Phe Val Thr Leu Arg Cys Thr Ala Ile Gly Ile Pro Val
225                 230                 235                 240

Pro Thr Ile Ser Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser
                245                 250                 255

Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu
            260                 265                 270

Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His
        275                 280                 285

Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Ala Thr Val Ser Ile Ala
    290                 295                 300

Glu Trp Ser Lys Ser Gln Lys Asp Ser Gln Gly Tyr Cys Ala Gln Tyr
305                 310                 315                 320

Arg Gly Glu Val Cys Asp Ala Val Leu Ala Lys Asp Ala Leu Val Phe
                325                 330                 335

Phe Asn Thr Ser Tyr Arg Asp Pro Glu Asp Ala Gln Glu Leu Leu Ile
            340                 345                 350

His Thr Ala Trp Asn Glu Leu Lys Ala Val Ser Pro Leu Cys Arg Pro
        355                 360                 365

Ala Ala Glu Ala Leu Leu Cys Asn His Leu Phe Gln Glu Cys Ser Pro
    370                 375                 380

Gly Val Val Pro Thr Pro Met Pro Ile Cys Arg Glu Tyr Cys Leu Ala
385                 390                 395                 400

Val Lys Glu Leu Phe Cys Ala Lys Glu Trp Gln Ala Met Glu Gly Lys
                405                 410                 415

Ala His Arg Gly Leu Tyr Arg Ser Gly Met His Leu Leu Pro Val Pro
            420                 425                 430

Glu Cys Ser Lys Leu Pro Ser Met His Arg Asp Pro Thr Ala Cys Thr
        435                 440                 445

Arg Leu Pro Tyr Leu Asp Tyr Lys Lys Glu Asn Ile Thr Thr Phe Pro
    450                 455                 460

Ser Ile Thr Ser Ser Arg Pro Ser Ala Asp Ile Pro Asn Leu
465                 470                 475
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACACTGCGTG GAATGAGCTG A 21

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AAATATGGCA GTCTTGTGCA 20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2580 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..2580

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
ATG AGA GAG CTT GTC AAC ATT CCA CTG TTA CAG ATG CTC ACC CTG GTT        48
Met Arg Glu Leu Val Asn Ile Pro Leu Leu Gln Met Leu Thr Leu Val
 1               5                  10                  15

GCC TTC AGC GGG ACT GAG AAA CTT CCA AAA GCC CCT GTC ATC ACC ACG        96
Ala Phe Ser Gly Thr Glu Lys Leu Pro Lys Ala Pro Val Ile Thr Thr
            20                  25                  30

CCT CTT GAA ACT GTA GAT GCC TTG GTT GAA GAA GTA GCG ACT TTC ATG       144
Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Val Ala Thr Phe Met
        35                  40                  45

TGT GCC GTG GAA TCC TAC CCT CAG CCC GAG ATT TCT TGG ACC AGA AAT       192
Cys Ala Val Glu Ser Tyr Pro Gln Pro Glu Ile Ser Trp Thr Arg Asn
    50                  55                  60

AAA ATT CTC ATT AAG CTG TTT GAC ACC CGC TAC AGC ATC CGG GAG AAT       240
Lys Ile Leu Ile Lys Leu Phe Asp Thr Arg Tyr Ser Ile Arg Glu Asn
65                  70                  75                  80
```

```
GGT CAG CTC CTC ACC ATT CTG AGC GTG GAA GAC AGT GAT GAT GGC ATC        288
Gly Gln Leu Leu Thr Ile Leu Ser Val Glu Asp Ser Asp Asp Gly Ile
                85                  90                  95

TAC TGC TGC ATA GCC AAC AAT GGA GTG GGA GGA GCC GTG GAG AGT TGT        336
Tyr Cys Cys Ile Ala Asn Asn Gly Val Gly Gly Ala Val Glu Ser Cys
            100                 105                 110

GGT GCC CTG CAA GTG AAG ATG AAA CCT AAA ATA ACT CGT CCT CCC ATT        384
Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Thr Arg Pro Pro Ile
        115                 120                 125

AAT GTA AAA ATA ATA GAG GGA TTG AAG GCA GTT CTG CCG TGC ACT ACG        432
Asn Val Lys Ile Ile Glu Gly Leu Lys Ala Val Leu Pro Cys Thr Thr
    130                 135                 140

ATG GGT AAC CCC AAA CCA TCT GTG TCC TGG ATC AAG GGG GAC AAT GCT        480
Met Gly Asn Pro Lys Pro Ser Val Ser Trp Ile Lys Gly Asp Asn Ala
145                 150                 155                 160

CTC AGG GAA AAT TCC AGA ATC GCA GTT CTT GAA TCT GGG AGC TTA AGG        528
Leu Arg Glu Asn Ser Arg Ile Ala Val Leu Glu Ser Gly Ser Leu Arg
                165                 170                 175

ATC CAT AAT GTG CAA AAG GAA GAT GCA GGA CAG TAC CGC TGT GTG GCC        576
Ile His Asn Val Gln Lys Glu Asp Ala Gly Gln Tyr Arg Cys Val Ala
            180                 185                 190

AAA AAC AGC CTG GGC ACA GCT TAC TCC AAA CTG GTG AAG CTG GAA GTG        624
Lys Asn Ser Leu Gly Thr Ala Tyr Ser Lys Leu Val Lys Leu Glu Val
        195                 200                 205

GAG GTT TTT GCA AGA ATC CTG CGT GCT CCT GAA TCC CAC AAT GTC ACC        672
Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr
    210                 215                 220

TTT GGT TCC TTT GTA ACC CTA CGC TGC ACA GCA ATA GGC ATC CCT GTC        720
Phe Gly Ser Phe Val Thr Leu Arg Cys Thr Ala Ile Gly Ile Pro Val
225                 230                 235                 240

CCC ACC ATC AGC TGG ATT GAA AAC GGA AAT GCT GTT TCT TCA GGT TCC        768
Pro Thr Ile Ser Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser
                245                 250                 255

ATT CAA GAG AGT GTG AAA GAC CGA GTG ATT GAC TCA AGA CTC CAG CTC        816
Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu
            260                 265                 270

TTC ATC ACA AAG CCA GGA CTC TAC ACA TGC ATA GCT ACC AAT AAG CAC        864
Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His
        275                 280                 285

GGA GAA AAG TTC AGT ACC GCA AAG GCT GCA GCC ACT GTC AGC ATA GCA        912
Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Ala Thr Val Ser Ile Ala
    290                 295                 300

GAA TGG AGT AAG TCA CAG AAA GAC AGC CAA GGC TAC TGT GCC CAG TAC        960
Glu Trp Ser Lys Ser Gln Lys Asp Ser Gln Gly Tyr Cys Ala Gln Tyr
305                 310                 315                 320
```

```
AGA GGG GAG GTG TGT GAT GCA GTC CTG GCG AAA GAT GCT CTT GTC TTC      1008
Arg Gly Glu Val Cys Asp Ala Val Leu Ala Lys Asp Ala Leu Val Phe
                325                 330                 335

TTC AAC ACC TCC TAC CGG GAC CCC GAG GAC GCC CAG GAG CTG CTG ATC      1056
Phe Asn Thr Ser Tyr Arg Asp Pro Glu Asp Ala Gln Glu Leu Leu Ile
            340                 345                 350

CAC ACT GCG TGG AAT GAG CTG AAG GCT GTG AGT CCA CTG TGC CGG CCA      1104
His Thr Ala Trp Asn Glu Leu Lys Ala Val Ser Pro Leu Cys Arg Pro
        355                 360                 365

GCT GCT GAG GCT CTG CTG TGT AAC CAC CTC TTC CAA GAG TGC AGC CCT      1152
Ala Ala Glu Ala Leu Leu Cys Asn His Leu Phe Gln Glu Cys Ser Pro
    370                 375                 380

GGA GTG GTA CCT ACT CCC ATG CCC ATT TGC AGA GAG TAC TGC CTG GCG      1200
Gly Val Val Pro Thr Pro Met Pro Ile Cys Arg Glu Tyr Cys Leu Ala
385                 390                 395                 400

GTA AAG GAG CTC TTC TGT GCA AAG GAA TGG CAG GCA ATG GAA GGA AAG      1248
Val Lys Glu Leu Phe Cys Ala Lys Glu Trp Gln Ala Met Glu Gly Lys
                405                 410                 415

GCC CAC CGG GGC CTC TAC AGA TCT GGG ATG CAT CTC CTT CCG GTA CCA      1296
Ala His Arg Gly Leu Tyr Arg Ser Gly Met His Leu Leu Pro Val Pro
            420                 425                 430

GAG TGC AGC AAG CTT CCC AGC ATG CAC CGG GAC CCC ACA GCC TGC ACA      1344
Glu Cys Ser Lys Leu Pro Ser Met His Arg Asp Pro Thr Ala Cys Thr
        435                 440                 445

AGA CTG CCA TAT TTA GCA TTC CCG TCA ATA ACG TCC TCC AGG CCG AGC      1392
Arg Leu Pro Tyr Leu Ala Phe Pro Ser Ile Thr Ser Ser Arg Pro Ser
    450                 455                 460

GCG GAC ATT CCA AAC CTG CCT GCC TCC ACC TCT TCC TTT GCC GTC TCG      1440
Ala Asp Ile Pro Asn Leu Pro Ala Ser Thr Ser Ser Phe Ala Val Ser
465                 470                 475                 480

CCT GCG TAC TCC ATG ACC GTC ATC ATC TCC ATC GTG TCC AGC TTT GCC      1488
Pro Ala Tyr Ser Met Thr Val Ile Ile Ser Ile Val Ser Ser Phe Ala
                485                 490                 495

CTG TTT GCT CTT CTC ACC ATC GCT ACT CTC TAT TGC TGC CGA AGG AGG      1536
Leu Phe Ala Leu Leu Thr Ile Ala Thr Leu Tyr Cys Cys Arg Arg Arg
            500                 505                 510

AAA GAA TGG AAA AAT AAG AAA AGA GAG TCG ACC GCG GTG ACC CTC ACC      1584
Lys Glu Trp Lys Asn Lys Lys Arg Glu Ser Thr Ala Val Thr Leu Thr
        515                 520                 525

ACG TTG CCT TCC GAG CTC CTG CTG GAT AGG CTC CAT CCC AAC CCC ATG      1632
Thr Leu Pro Ser Glu Leu Leu Leu Asp Arg Leu His Pro Asn Pro Met
    530                 535                 540

TAC CAG AGG ATG CCA CTC CTT CTG AAT CCT AAG TTG CTC AGC CTG GAG      1680
Tyr Gln Arg Met Pro Leu Leu Leu Asn Pro Lys Leu Leu Ser Leu Glu
545                 550                 555                 560
```

```
TAT CCG AGG AAT AAC ATT GAG TAT GTC CGA GAC ATC GGA GAG GGG GCG      1728
Tyr Pro Arg Asn Asn Ile Glu Tyr Val Arg Asp Ile Gly Glu Gly Ala
                565                 570                 575

TTT GGA AGA GTC TTC CAA GCA AGG GCC CCT GGC TTG CTG CCT TAT GAA      1776
Phe Gly Arg Val Phe Gln Ala Arg Ala Pro Gly Leu Leu Pro Tyr Glu
            580                 585                 590

CCT TTC ACT ATG GTG GCC GTG AAG ATG CTT AAG GAA GAG GCC TCT GCA      1824
Pro Phe Thr Met Val Ala Val Lys Met Leu Lys Glu Glu Ala Ser Ala
        595                 600                 605

GAC ATG CAA GCG GAC TTT CAG AGG GAG GCG GCC CTC ATG GCA GAG TTT      1872
Asp Met Gln Ala Asp Phe Gln Arg Glu Ala Ala Leu Met Ala Glu Phe
    610                 615                 620

GAC AAC CCC AAC ATC GTG AAA CTC TTA GGT GTG TGT GCC GTT GGG AAG      1920
Asp Asn Pro Asn Ile Val Lys Leu Leu Gly Val Cys Ala Val Gly Lys
625                 630                 635                 640

CCG ATG TGT CTG CTC TTT GAA TAT ATG GCC TAT GGT GAC CTC AAT GAG      1968
Pro Met Cys Leu Leu Phe Glu Tyr Met Ala Tyr Gly Asp Leu Asn Glu
                645                 650                 655

TTC CTC CGA AGT ATG TCC CCG CAC ACT GTT TGC AGC CTC AGC CAC AGT      2016
Phe Leu Arg Ser Met Ser Pro His Thr Val Cys Ser Leu Ser His Ser
            660                 665                 670

GAC CTG TCC ACG AGG GCT CGG GTG TCT AGC CCT GGT CCT CCA CCA CTG      2064
Asp Leu Ser Thr Arg Ala Arg Val Ser Ser Pro Gly Pro Pro Pro Leu
        675                 680                 685

TCC TGT GCA GAA CAG CTC TGC ATT GCC AGG CAG GTG GCA GCT GGC ATG      2112
Ser Cys Ala Glu Gln Leu Cys Ile Ala Arg Gln Val Ala Ala Gly Met
    690                 695                 700

GCC TAC CTT TCA GAG CGC AAG TTT GTC CAC CGG GAC TTA GCT ACC AGG      2160
Ala Tyr Leu Ser Glu Arg Lys Phe Val His Arg Asp Leu Ala Thr Arg
705                 710                 715                 720

AAC TGC CTG GTT GGG GAG ACC ATG GTG GTG AAA ATT GCA GAC TTT GGC      2208
Asn Cys Leu Val Gly Glu Thr Met Val Val Lys Ile Ala Asp Phe Gly
                725                 730                 735

CTC TCC AGG AAC ATC TAT TCC GCA GAC TAC TAC AAA GCT GAT GGA AAT      2256
Leu Ser Arg Asn Ile Tyr Ser Ala Asp Tyr Tyr Lys Ala Asp Gly Asn
            740                 745                 750

GAC GCC ATC CCT ATC CGC TGG ATG CCG CCC GAG TCT ATC TTC TAC AAC      2304
Asp Ala Ile Pro Ile Arg Trp Met Pro Pro Glu Ser Ile Phe Tyr Asn
        755                 760                 765

CGC TAC ACC ACG GAG TCG GAT GTA TGG GCC TAT GGT GTG GTC CTC TGG      2352
Arg Tyr Thr Thr Glu Ser Asp Val Trp Ala Tyr Gly Val Val Leu Trp
    770                 775                 780

GAG ATC TTC TCC TAT GGG CTG CAG CCC TAC TAT GGA ATG GCC CAC GAG      2400
Glu Ile Phe Ser Tyr Gly Leu Gln Pro Tyr Tyr Gly Met Ala His Glu
785                 790                 795                 800
```

```
GAG GTC ATT TAC TAT GTG AGA GAT GGC AAC ATC CTC GCC TGC CCT GAG      2448
Glu Val Ile Tyr Tyr Val Arg Asp Gly Asn Ile Leu Ala Cys Pro Glu
                805                 810                 815

AAC TGC CCC TTG GAA CTG TAC AAC CTC ATG CGC CTG TGT TGG AGC AAG      2496
Asn Cys Pro Leu Glu Leu Tyr Asn Leu Met Arg Leu Cys Trp Ser Lys
            820                 825                 830

CTG CCT GCT GAT AGA CCC AGC TTC TGC AGT ATC CAC AGG ATC CTG CAG      2544
Leu Pro Ala Asp Arg Pro Ser Phe Cys Ser Ile His Arg Ile Leu Gln
        835                 840                 845

CGC ATG TGC GAG AGA GCA GAG GGA ACG GTG GGT GTC                      2580
Arg Met Cys Glu Arg Ala Glu Gly Thr Val Gly Val
    850                 855                 860
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 860 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Arg Glu Leu Val Asn Ile Pro Leu Leu Gln Met Leu Thr Leu Val
 1               5                  10                  15

Ala Phe Ser Gly Thr Glu Lys Leu Pro Lys Ala Pro Val Ile Thr Thr
            20                  25                  30

Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Val Ala Thr Phe Met
        35                  40                  45

Cys Ala Val Glu Ser Tyr Pro Gln Pro Glu Ile Ser Trp Thr Arg Asn
    50                  55                  60

Lys Ile Leu Ile Lys Leu Phe Asp Thr Arg Tyr Ser Ile Arg Glu Asn
65                  70                  75                  80

Gly Gln Leu Leu Thr Ile Leu Ser Val Glu Asp Ser Asp Asp Gly Ile
                85                  90                  95

Tyr Cys Cys Ile Ala Asn Asn Gly Val Gly Gly Ala Val Glu Ser Cys
            100                 105                 110

Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Thr Arg Pro Pro Ile
        115                 120                 125

Asn Val Lys Ile Ile Glu Gly Leu Lys Ala Val Leu Pro Cys Thr Thr
    130                 135                 140

Met Gly Asn Pro Lys Pro Ser Val Ser Trp Ile Lys Gly Asp Asn Ala
145                 150                 155                 160

Leu Arg Glu Asn Ser Arg Ile Ala Val Leu Glu Ser Gly Ser Leu Arg
                165                 170                 175
```

```
Ile His Asn Val Gln Lys Glu Asp Ala Gly Gln Tyr Arg Cys Val Ala
            180                 185                 190

Lys Asn Ser Leu Gly Thr Ala Tyr Ser Lys Leu Val Lys Leu Glu Val
        195                 200                 205

Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr
    210                 215                 220

Phe Gly Ser Phe Val Thr Leu Arg Cys Thr Ala Ile Gly Ile Pro Val
225                 230                 235                 240

Pro Thr Ile Ser Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser
                245                 250                 255

Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu
            260                 265                 270

Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His
        275                 280                 285

Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Ala Thr Val Ser Ile Ala
    290                 295                 300

Glu Trp Ser Lys Ser Gln Lys Asp Ser Gln Gly Tyr Cys Ala Gln Tyr
305                 310                 315                 320

Arg Gly Glu Val Cys Asp Ala Val Leu Ala Lys Asp Ala Leu Val Phe
                325                 330                 335

Phe Asn Thr Ser Tyr Arg Asp Pro Glu Asp Ala Gln Glu Leu Leu Ile
            340                 345                 350

His Thr Ala Trp Asn Glu Leu Lys Ala Val Ser Pro Leu Cys Arg Pro
        355                 360                 365

Ala Ala Glu Ala Leu Leu Cys Asn His Leu Phe Gln Glu Cys Ser Pro
    370                 375                 380

Gly Val Val Pro Thr Pro Met Pro Ile Cys Arg Glu Tyr Cys Leu Ala
385                 390                 395                 400

Val Lys Glu Leu Phe Cys Ala Lys Glu Trp Gln Ala Met Glu Gly Lys
                405                 410                 415

Ala His Arg Gly Leu Tyr Arg Ser Gly Met His Leu Leu Pro Val Pro
            420                 425                 430

Glu Cys Ser Lys Leu Pro Ser Met His Arg Asp Pro Thr Ala Cys Thr
        435                 440                 445

Arg Leu Pro Tyr Leu Ala Phe Pro Ser Ile Thr Ser Ser Arg Pro Ser
    450                 455                 460

Ala Asp Ile Pro Asn Leu Pro Ala Ser Thr Ser Ser Phe Ala Val Ser
465                 470                 475                 480

Pro Ala Tyr Ser Met Thr Val Ile Ile Ser Ile Val Ser Ser Phe Ala
                485                 490                 495

Leu Phe Ala Leu Leu Thr Ile Ala Thr Leu Tyr Cys Cys Arg Arg Arg
            500                 505                 510
```

Lys Glu Trp Lys Asn Lys Lys Arg Glu Ser Thr Ala Val Thr Leu Thr
515 520 525

Thr Leu Pro Ser Glu Leu Leu Leu Asp Arg Leu His Pro Asn Pro Met
530 535 540

Tyr Gln Arg Met Pro Leu Leu Leu Asn Pro Lys Leu Leu Ser Leu Glu
545 550 555 560

Tyr Pro Arg Asn Asn Ile Glu Tyr Val Arg Asp Ile Gly Glu Gly Ala
565 570 575

Phe Gly Arg Val Phe Gln Ala Arg Ala Pro Gly Leu Leu Pro Tyr Glu
580 585 590

Pro Phe Thr Met Val Ala Val Lys Met Leu Lys Glu Glu Ala Ser Ala
595 600 605

Asp Met Gln Ala Asp Phe Gln Arg Glu Ala Ala Leu Met Ala Glu Phe
610 615 620

Asp Asn Pro Asn Ile Val Lys Leu Leu Gly Val Cys Ala Val Gly Lys
625 630 635 640

Pro Met Cys Leu Leu Phe Glu Tyr Met Ala Tyr Gly Asp Leu Asn Glu
645 650 655

Phe Leu Arg Ser Met Ser Pro His Thr Val Cys Ser Leu Ser His Ser
660 665 670

Asp Leu Ser Thr Arg Ala Arg Val Ser Ser Pro Gly Pro Pro Pro Leu
675 680 685

Ser Cys Ala Glu Gln Leu Cys Ile Ala Arg Gln Val Ala Ala Gly Met
690 695 700

Ala Tyr Leu Ser Glu Arg Lys Phe Val His Arg Asp Leu Ala Thr Arg
705 710 715 720

Asn Cys Leu Val Gly Glu Thr Met Val Val Lys Ile Ala Asp Phe Gly
725 730 735

Leu Ser Arg Asn Ile Tyr Ser Ala Asp Tyr Tyr Lys Ala Asp Gly Asn
740 745 750

Asp Ala Ile Pro Ile Arg Trp Met Pro Pro Glu Ser Ile Phe Tyr Asn
755 760 765

Arg Tyr Thr Thr Glu Ser Asp Val Trp Ala Tyr Gly Val Val Leu Trp
770 775 780

Glu Ile Phe Ser Tyr Gly Leu Gln Pro Tyr Tyr Gly Met Ala His Glu
785 790 795 800

Glu Val Ile Tyr Tyr Val Arg Asp Gly Asn Ile Leu Ala Cys Pro Glu
805 810 815

Asn Cys Pro Leu Glu Leu Tyr Asn Leu Met Arg Leu Cys Trp Ser Lys
820 825 830

Leu Pro Ala Asp Arg Pro Ser Phe Cys Ser Ile His Arg Ile Leu Gln
835 840 845

```
Arg Met Cys Glu Arg Ala Glu Gly Thr Val Gly Val
    850                 855                 860
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2604 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..2604

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
ATG AGA GAG CTT GTC AAC ATT CCA CTG TTA CAG ATG CTC ACC CTG GTT        48
Met Arg Glu Leu Val Asn Ile Pro Leu Leu Gln Met Leu Thr Leu Val
  1               5                  10                  15

GCC TTC AGC GGG ACT GAG AAA CTT CCA AAA GCC CCT GTC ATC ACC ACG        96
Ala Phe Ser Gly Thr Glu Lys Leu Pro Lys Ala Pro Val Ile Thr Thr
            20                  25                  30

CCT CTT GAA ACT GTA GAT GCC TTG GTT GAA GAA GTA GCG ACT TTC ATG       144
Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Val Ala Thr Phe Met
        35                  40                  45

TGT GCC GTG GAA TCC TAC CCT CAG CCC GAG ATT TCT TGG ACC AGA AAT       192
Cys Ala Val Glu Ser Tyr Pro Gln Pro Glu Ile Ser Trp Thr Arg Asn
    50                  55                  60

AAA ATT CTC ATT AAG CTG TTT GAC ACC CGC TAC AGC ATC CGG GAG AAT       240
Lys Ile Leu Ile Lys Leu Phe Asp Thr Arg Tyr Ser Ile Arg Glu Asn
65                  70                  75                  80

GGT CAG CTC CTC ACC ATT CTG AGC GTG GAA GAC AGT GAT GAT GGC ATC       288
Gly Gln Leu Leu Thr Ile Leu Ser Val Glu Asp Ser Asp Asp Gly Ile
                85                  90                  95

TAC TGC TGC ATA GCC AAC AAT GGA GTG GGA GGA GCC GTG GAG AGT TGT       336
Tyr Cys Cys Ile Ala Asn Asn Gly Val Gly Gly Ala Val Glu Ser Cys
            100                 105                 110

GGT GCC CTG CAA GTG AAG ATG AAA CCT AAA ATA ACT CGT CCT CCC ATT       384
Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Thr Arg Pro Pro Ile
        115                 120                 125

AAT GTA AAA ATA ATA GAG GGA TTG AAG GCA GTT CTG CCG TGC ACT ACG       432
Asn Val Lys Ile Ile Glu Gly Leu Lys Ala Val Leu Pro Cys Thr Thr
    130                 135                 140

ATG GGT AAC CCC AAA CCA TCT GTG TCC TGG ATC AAG GGG GAC AAT GCT       480
Met Gly Asn Pro Lys Pro Ser Val Ser Trp Ile Lys Gly Asp Asn Ala
145                 150                 155                 160
```

```
CTC AGG GAA AAT TCC AGA ATC GCA GTT CTT GAA TCT GGG AGC TTA AGG      528
Leu Arg Glu Asn Ser Arg Ile Ala Val Leu Glu Ser Gly Ser Leu Arg
                165                 170                 175

ATC CAT AAT GTG CAA AAG GAA GAT GCA GGA CAG TAC CGC TGT GTG GCC      576
Ile His Asn Val Gln Lys Glu Asp Ala Gly Gln Tyr Arg Cys Val Ala
            180                 185                 190

AAA AAC AGC CTG GGC ACA GCT TAC TCC AAA CTG GTG AAG CTG GAA GTG      624
Lys Asn Ser Leu Gly Thr Ala Tyr Ser Lys Leu Val Lys Leu Glu Val
        195                 200                 205

GAG GTT TTT GCA AGA ATC CTG CGT GCT CCT GAA TCC CAC AAT GTC ACC      672
Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr
    210                 215                 220

TTT GGT TCC TTT GTA ACC CTA CGC TGC ACA GCA ATA GGC ATC CCT GTC      720
Phe Gly Ser Phe Val Thr Leu Arg Cys Thr Ala Ile Gly Ile Pro Val
225                 230                 235                 240

CCC ACC ATC AGC TGG ATT GAA AAC GGA AAT GCT GTT TCT TCA GGT TCC      768
Pro Thr Ile Ser Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser
                245                 250                 255

ATT CAA GAG AGT GTG AAA GAC CGA GTG ATT GAC TCA AGA CTC CAG CTC      816
Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu
            260                 265                 270

TTC ATC ACA AAG CCA GGA CTC TAC ACA TGC ATA GCT ACC AAT AAG CAC      864
Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His
        275                 280                 285

GGA GAA AAG TTC AGT ACC GCA AAG GCT GCA GCC ACT GTC AGC ATA GCA      912
Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Ala Thr Val Ser Ile Ala
    290                 295                 300

GAA TGG AGT AAG TCA CAG AAA GAC AGC CAA GGC TAC TGT GCC CAG TAC      960
Glu Trp Ser Lys Ser Gln Lys Asp Ser Gln Gly Tyr Cys Ala Gln Tyr
305                 310                 315                 320

AGA GGG GAG GTG TGT GAT GCA GTC CTG GCG AAA GAT GCT CTT GTC TTC     1008
Arg Gly Glu Val Cys Asp Ala Val Leu Ala Lys Asp Ala Leu Val Phe
                325                 330                 335

TTC AAC ACC TCC TAC CGG GAC CCC GAG GAC GCC CAG GAG CTG CTG ATC     1056
Phe Asn Thr Ser Tyr Arg Asp Pro Glu Asp Ala Gln Glu Leu Leu Ile
            340                 345                 350

CAC ACT GCG TGG AAT GAG CTG AAG GCT GTG AGT CCA CTG TGC CGG CCA     1104
His Thr Ala Trp Asn Glu Leu Lys Ala Val Ser Pro Leu Cys Arg Pro
        355                 360                 365

GCT GCT GAG GCT CTG CTG TGT AAC CAC CTC TTC CAA GAG TGC AGC CCT     1152
Ala Ala Glu Ala Leu Leu Cys Asn His Leu Phe Gln Glu Cys Ser Pro
    370                 375                 380

GGA GTG GTA CCT ACT CCC ATG CCC ATT TGC AGA GAG TAC TGC CTG GCG     1200
Gly Val Val Pro Thr Pro Met Pro Ile Cys Arg Glu Tyr Cys Leu Ala
385                 390                 395                 400
```

```
GTA AAG GAG CTC TTC TGT GCA AAG GAA TGG CAG GCA ATG GAA GGA AAG      1248
Val Lys Glu Leu Phe Cys Ala Lys Glu Trp Gln Ala Met Glu Gly Lys
                405                 410                 415

GCC CAC CGG GGC CTC TAC AGA TCT GGG ATG CAT CTC CTT CCG GTA CCA      1296
Ala His Arg Gly Leu Tyr Arg Ser Gly Met His Leu Leu Pro Val Pro
            420                 425                 430

GAG TGC AGC AAG CTT CCC AGC ATG CAC CGG GAC CCC ACA GCC TGC ACA      1344
Glu Cys Ser Lys Leu Pro Ser Met His Arg Asp Pro Thr Ala Cys Thr
        435                 440                 445

AGA CTG CCA TAT TTA GAT TAT AAA AAA GAA AAC ATA ACA ACA TTC CCG      1392
Arg Leu Pro Tyr Leu Asp Tyr Lys Lys Glu Asn Ile Thr Thr Phe Pro
    450                 455                 460

TCA ATA ACG TCC TCC AGG CCG AGC GCG GAC ATT CCA AAC CTG CCT GCC      1440
Ser Ile Thr Ser Ser Arg Pro Ser Ala Asp Ile Pro Asn Leu Pro Ala
465                 470                 475                 480

TCC ACC TCT TCC TTT GCC GTC TCG CCT GCG TAC TCC ATG ACC GTC ATC      1488
Ser Thr Ser Ser Phe Ala Val Ser Pro Ala Tyr Ser Met Thr Val Ile
                485                 490                 495

ATC TCC ATC GTG TCC AGC TTT GCC CTG TTT GCT CTT CTC ACC ATC GCT      1536
Ile Ser Ile Val Ser Ser Phe Ala Leu Phe Ala Leu Leu Thr Ile Ala
            500                 505                 510

ACT CTC TAT TGC TGC CGA AGG AGG AAA GAA TGG AAA AAT AAG AAA AGA      1584
Thr Leu Tyr Cys Cys Arg Arg Arg Lys Glu Trp Lys Asn Lys Lys Arg
        515                 520                 525

GAG TCG ACC GCG GTG ACC CTC ACC ACG TTG CCT TCC GAG CTC CTG CTG      1632
Glu Ser Thr Ala Val Thr Leu Thr Thr Leu Pro Ser Glu Leu Leu Leu
    530                 535                 540

GAT AGG CTC CAT CCC AAC CCC ATG TAC CAG AGG ATG CCA CTC CTT CTG      1680
Asp Arg Leu His Pro Asn Pro Met Tyr Gln Arg Met Pro Leu Leu Leu
545                 550                 555                 560

AAT CCT AAG TTG CTC AGC CTG GAG TAT CCG AGG AAT AAC ATT GAG TAT      1728
Asn Pro Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile Glu Tyr
                565                 570                 575

GTC CGA GAC ATC GGA GAG GGG GCG TTT GGA AGA GTC TTC CAA GCA AGG      1776
Val Arg Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln Ala Arg
            580                 585                 590

GCC CCT GGC TTG CTG CCT TAT GAA CCT TTC ACT ATG GTG GCC GTG AAG      1824
Ala Pro Gly Leu Leu Pro Tyr Glu Pro Phe Thr Met Val Ala Val Lys
        595                 600                 605

ATG CTT AAG GAA GAG GCC TCT GCA GAC ATG CAA GCG GAC TTT CAG AGG      1872
Met Leu Lys Glu Glu Ala Ser Ala Asp Met Gln Ala Asp Phe Gln Arg
    610                 615                 620

GAG GCG GCC CTC ATG GCA GAG TTT GAC AAC CCC AAC ATC GTG AAA CTC      1920
Glu Ala Ala Leu Met Ala Glu Phe Asp Asn Pro Asn Ile Val Lys Leu
625                 630                 635                 640
```

```
TTA GGT GTG TGT GCC GTT GGG AAG CCG ATG TGT CTG CTC TTT GAA TAT      1968
Leu Gly Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe Glu Tyr
                645                 650                 655

ATG GCC TAT GGT GAC CTC AAT GAG TTC CTC CGA AGT ATG TCC CCG CAC      2016
Met Ala Tyr Gly Asp Leu Asn Glu Phe Leu Arg Ser Met Ser Pro His
            660                 665                 670

ACT GTT TGC AGC CTC AGC CAC AGT GAC CTG TCC ACG AGG GCT CGG GTG      2064
Thr Val Cys Ser Leu Ser His Ser Asp Leu Ser Thr Arg Ala Arg Val
        675                 680                 685

TCT AGC CCT GGT CCT CCA CCA CTG TCC TGT GCA GAA CAG CTC TGC ATT      2112
Ser Ser Pro Gly Pro Pro Pro Leu Ser Cys Ala Glu Gln Leu Cys Ile
    690                 695                 700

GCC AGG CAG GTG GCA GCT GGC ATG GCC TAC CTT TCA GAG CGC AAG TTT      2160
Ala Arg Gln Val Ala Ala Gly Met Ala Tyr Leu Ser Glu Arg Lys Phe
705                 710                 715                 720

GTC CAC CGG GAC TTA GCT ACC AGG AAC TGC CTG GTT GGG GAG ACC ATG      2208
Val His Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Thr Met
                725                 730                 735

GTG GTG AAA ATT GCA GAC TTT GGC CTC TCC AGG AAC ATC TAT TCC GCA      2256
Val Val Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr Ser Ala
            740                 745                 750

GAC TAC TAC AAA GCT GAT GGA AAT GAC GCC ATC CCT ATC CGC TGG ATG      2304
Asp Tyr Tyr Lys Ala Asp Gly Asn Asp Ala Ile Pro Ile Arg Trp Met
        755                 760                 765

CCG CCC GAG TCT ATC TTC TAC AAC CGC TAC ACC ACG GAG TCG GAT GTA      2352
Pro Pro Glu Ser Ile Phe Tyr Asn Arg Tyr Thr Thr Glu Ser Asp Val
    770                 775                 780

TGG GCC TAT GGT GTG GTC CTC TGG GAG ATC TTC TCC TAT GGG CTG CAG      2400
Trp Ala Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Tyr Gly Leu Gln
785                 790                 795                 800

CCC TAC TAT GGA ATG GCC CAC GAG GAG GTC ATT TAC TAT GTG AGA GAT      2448
Pro Tyr Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg Asp
                805                 810                 815

GGC AAC ATC CTC GCC TGC CCT GAG AAC TGC CCC TTG GAA CTG TAC AAC      2496
Gly Asn Ile Leu Ala Cys Pro Glu Asn Cys Pro Leu Glu Leu Tyr Asn
            820                 825                 830

CTC ATG CGC CTG TGT TGG AGC AAG CTG CCT GCT GAT AGA CCC AGC TTC      2544
Leu Met Arg Leu Cys Trp Ser Lys Leu Pro Ala Asp Arg Pro Ser Phe
        835                 840                 845

TGC AGT ATC CAC AGG ATC CTG CAG CGC ATG TGC GAG AGA GCA GAG GGA      2592
Cys Ser Ile His Arg Ile Leu Gln Arg Met Cys Glu Arg Ala Glu Gly
    850                 855                 860

ACG GTG GGT GTC                                                     2604
Thr Val Gly Val
865
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 868 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Met Arg Glu Leu Val Asn Ile Pro Leu Leu Gln Met Leu Thr Leu Val
 1               5                  10                  15

Ala Phe Ser Gly Thr Glu Lys Leu Pro Lys Ala Pro Val Ile Thr Thr
            20                  25                  30

Pro Leu Glu Thr Val Asp Ala Leu Val Glu Glu Val Ala Thr Phe Met
        35                  40                  45

Cys Ala Val Glu Ser Tyr Pro Gln Pro Glu Ile Ser Trp Thr Arg Asn
    50                  55                  60

Lys Ile Leu Ile Lys Leu Phe Asp Thr Arg Tyr Ser Ile Arg Glu Asn
65                  70                  75                  80

Gly Gln Leu Leu Thr Ile Leu Ser Val Glu Asp Ser Asp Asp Gly Ile
                85                  90                  95

Tyr Cys Cys Ile Ala Asn Asn Gly Val Gly Gly Ala Val Glu Ser Cys
            100                 105                 110

Gly Ala Leu Gln Val Lys Met Lys Pro Lys Ile Thr Arg Pro Pro Ile
        115                 120                 125

Asn Val Lys Ile Ile Glu Gly Leu Lys Ala Val Leu Pro Cys Thr Thr
    130                 135                 140

Met Gly Asn Pro Lys Pro Ser Val Ser Trp Ile Lys Gly Asp Asn Ala
145                 150                 155                 160

Leu Arg Glu Asn Ser Arg Ile Ala Val Leu Glu Ser Gly Ser Leu Arg
                165                 170                 175

Ile His Asn Val Gln Lys Glu Asp Ala Gly Gln Tyr Arg Cys Val Ala
            180                 185                 190

Lys Asn Ser Leu Gly Thr Ala Tyr Ser Lys Leu Val Lys Leu Glu Val
        195                 200                 205

Glu Val Phe Ala Arg Ile Leu Arg Ala Pro Glu Ser His Asn Val Thr
    210                 215                 220

Phe Gly Ser Phe Val Thr Leu Arg Cys Thr Ala Ile Gly Ile Pro Val
225                 230                 235                 240

Pro Thr Ile Ser Trp Ile Glu Asn Gly Asn Ala Val Ser Ser Gly Ser
                245                 250                 255

Ile Gln Glu Ser Val Lys Asp Arg Val Ile Asp Ser Arg Leu Gln Leu
            260                 265                 270
```

```
Phe Ile Thr Lys Pro Gly Leu Tyr Thr Cys Ile Ala Thr Asn Lys His
        275                 280                 285

Gly Glu Lys Phe Ser Thr Ala Lys Ala Ala Ala Thr Val Ser Ile Ala
    290                 295                 300

Glu Trp Ser Lys Ser Gln Lys Asp Ser Gln Gly Tyr Cys Ala Gln Tyr
305                 310                 315                 320

Arg Gly Glu Val Cys Asp Ala Val Leu Ala Lys Asp Ala Leu Val Phe
                325                 330                 335

Phe Asn Thr Ser Tyr Arg Asp Pro Glu Asp Ala Gln Glu Leu Leu Ile
            340                 345                 350

His Thr Ala Trp Asn Glu Leu Lys Ala Val Ser Pro Leu Cys Arg Pro
        355                 360                 365

Ala Ala Glu Ala Leu Leu Cys Asn His Leu Phe Gln Glu Cys Ser Pro
    370                 375                 380

Gly Val Val Pro Thr Pro Met Pro Ile Cys Arg Glu Tyr Cys Leu Ala
385                 390                 395                 400

Val Lys Glu Leu Phe Cys Ala Lys Glu Trp Gln Ala Met Glu Gly Lys
                405                 410                 415

Ala His Arg Gly Leu Tyr Arg Ser Gly Met His Leu Leu Pro Val Pro
            420                 425                 430

Glu Cys Ser Lys Leu Pro Ser Met His Arg Asp Pro Thr Ala Cys Thr
        435                 440                 445

Arg Leu Pro Tyr Leu Asp Tyr Lys Lys Glu Asn Ile Thr Thr Phe Pro
    450                 455                 460

Ser Ile Thr Ser Ser Arg Pro Ser Ala Asp Ile Pro Asn Leu Pro Ala
465                 470                 475                 480

Ser Thr Ser Ser Phe Ala Val Ser Pro Ala Tyr Ser Met Thr Val Ile
                485                 490                 495

Ile Ser Ile Val Ser Ser Phe Ala Leu Phe Ala Leu Leu Thr Ile Ala
            500                 505                 510

Thr Leu Tyr Cys Cys Arg Arg Arg Lys Glu Trp Lys Asn Lys Lys Arg
        515                 520                 525

Glu Ser Thr Ala Val Thr Leu Thr Thr Leu Pro Ser Glu Leu Leu Leu
    530                 535                 540

Asp Arg Leu His Pro Asn Pro Met Tyr Gln Arg Met Pro Leu Leu Leu
545                 550                 555                 560

Asn Pro Lys Leu Leu Ser Leu Glu Tyr Pro Arg Asn Asn Ile Glu Tyr
                565                 570                 575

Val Arg Asp Ile Gly Glu Gly Ala Phe Gly Arg Val Phe Gln Ala Arg
            580                 585                 590

Ala Pro Gly Leu Leu Pro Tyr Glu Pro Phe Thr Met Val Ala Val Lys
        595                 600                 605
```

```
Met Leu Lys Glu Glu Ala Ser Ala Asp Met Gln Ala Asp Phe Gln Arg
    610                 615                 620

Glu Ala Ala Leu Met Ala Glu Phe Asp Asn Pro Asn Ile Val Lys Leu
625                 630                 635                 640

Leu Gly Val Cys Ala Val Gly Lys Pro Met Cys Leu Leu Phe Glu Tyr
                645                 650                 655

Met Ala Tyr Gly Asp Leu Asn Glu Phe Leu Arg Ser Met Ser Pro His
            660                 665                 670

Thr Val Cys Ser Leu Ser His Ser Asp Leu Ser Thr Arg Ala Arg Val
        675                 680                 685

Ser Ser Pro Gly Pro Pro Pro Leu Ser Cys Ala Glu Gln Leu Cys Ile
    690                 695                 700

Ala Arg Gln Val Ala Ala Gly Met Ala Tyr Leu Ser Glu Arg Lys Phe
705                 710                 715                 720

Val His Arg Asp Leu Ala Thr Arg Asn Cys Leu Val Gly Glu Thr Met
                725                 730                 735

Val Val Lys Ile Ala Asp Phe Gly Leu Ser Arg Asn Ile Tyr Ser Ala
            740                 745                 750

Asp Tyr Tyr Lys Ala Asp Gly Asn Asp Ala Ile Pro Ile Arg Trp Met
        755                 760                 765

Pro Pro Glu Ser Ile Phe Tyr Asn Arg Tyr Thr Thr Glu Ser Asp Val
    770                 775                 780

Trp Ala Tyr Gly Val Val Leu Trp Glu Ile Phe Ser Tyr Gly Leu Gln
785                 790                 795                 800

Pro Tyr Tyr Gly Met Ala His Glu Glu Val Ile Tyr Tyr Val Arg Asp
                805                 810                 815

Gly Asn Ile Leu Ala Cys Pro Glu Asn Cys Pro Leu Glu Leu Tyr Asn
            820                 825                 830

Leu Met Arg Leu Cys Trp Ser Lys Leu Pro Ala Asp Arg Pro Ser Phe
        835                 840                 845

Cys Ser Ile His Arg Ile Leu Gln Arg Met Cys Glu Arg Ala Glu Gly
    850                 855                 860

Thr Val Gly Val
865
```

What is claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) the nucleotide sequence of SEQ ID NO:1 from nucleotide 13 to nucleotide 1602;

(b) the nucleotide sequence of SEQ ID NO:18 from nucleotide 1 to nucleotide 2580;

(c) the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 2604;

(d) a nucleotide sequence capable of hybridizing to a nucleic acid sequence specified in (a)-(c);

(e) a nucleotide sequence varying from the sequence of a nucleotide sequence specified in (a)-(c) as a result of degeneracy of the genetic code; and

(f) an allelic variant of a nucleotide sequence specified in (a)-(c).

2. The polynucleotide of claim 1 wherein said nucleotide sequence encodes for a protein having an activity selected from the group consisting of mlk receptor activity and mlk ligand binding activity.

3. The polynucleotide of claim 1 wherein said nucleotide sequence is operably linked to an expression control sequence.

4. The polynucleotide of claim 1 which encodes a murine mlk protein.

5. The polynucleotide of claim 1 which encodes a human mlk protein.

6. The polynucleotide of claim 1 comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 13 to nucleotide 1602.

7. The polynucleotide of claim 1 encoding the protein of claim 14.

8. The polynucleotide of claim 1 comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 1 to nucleotide 2580.

9. The polynulceotide of claim 1 comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 2604.

10. The polynucleotide of claim 2 which encodes a protein having mlk ligand binding activity.

11. A host cell transformed with the polynucleotide of claim 3.

12. The host cell of claim 11, wherein said cell is a mammalian cell.

13. A process for producing a mlk protein, said process comprising:

(a) growing a culture of the host cell of claim 11 in a suitable culture medium; and

(b) purifying the mlk protein from the culture.

14. An isolated mlk receptor protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) the amino acid sequence of SEQ ID NO:2 from amino acids 1 to 156;

(c) the amino acid sequence of SEQ ID NO:2 from amino acids 157 to 177;

(d) the amino acid sequence of SEQ ID NO:2 from amino acids 178 to 530;

(e) the amino acid sequence of SEQ ID NO:2 from amino acids 242 to 517;

(f) the amino acid sequence of SEQ ID NO:15;

(g) the amino acid sequence of SEQ ID NO:15 from amino acids 22 to 478;

(h) the amino acid sequence of SEQ ID NO:19;

(i) the amino acid sequence of SEQ ID NO:19 from amino acids 22 to 486;

(j) the amino acid sequence of SEQ ID NO:19 from amino acids 487 to 507;

(k) the amino acid sequence of SEQ ID NO:19 from amino acids 508 to 860;

(l) the amino acid sequence of SEQ ID NO:19 from amino acids 572 to 847;

(m) the amino acid sequence of SEQ ID NO:21;

(n) the amino acid sequence of SEQ ID NO:21 from amino acids 22 to 494;

(o) the amino acid sequence of SEQ ID NO:21 from amino acids 495 to 515;

(p) the amino acid sequence of SEQ ID NO:21 from amino acids 516 to 868;

(q) the amino acid sequence of SEQ ID NO:21 from amino acids 580 to 855;

(r) fragments of (a)-(q) having mlk receptor activity; and

(s) fragments of (a)-(q) having mlk ligand binding activity.

15. The protein of claim 14 comprising the amino acid sequence of SEQ ID NO:2.

16. The protein of claim 14 comprising the amino acid sequence of SEQ ID NO:15.

17. The protein of claim 14 comprising the amino acid sequence of SEQ ID NO:19.

18. The protein of claim 14 comprising the amino acid sequence of SEQ ID NO:21.

19. The protein of claim 14 comprising a fragment having mlk ligand binding activity.

20. A pharmaceutical composition comprising a mlk protein of claim 14 and a pharmaceutically acceptable carrier.

21. A mlk protein produced according to the process of claim 13.

22. A composition comprising an antibody which specifically reacts with a mlk protein of claim 14.

23. A method of identifying a mlk receptor ligand, said method comprising:

(a) providing a sample containing a potential source of mlk ligand;

(b) contacting said sample with a protein of claim 14; and

(c) collecting materials binding to said protein.

24. A mlk receptor ligand.

25. The ligand of claim 24 identified according to the method of claim 23.

26. A pharmaceutical composition comprising a ligand of claim 24 and a pharmaceutically acceptable carrier.

27. An isolated polynucleotide encoding a ligand of claim 24.

28. A host cell transformed with an expression vector comprising the polynucleotide of claim 27.

29. A method of treating a mlk-related condition, said method comprising administering a therapeutically effective amount of a composition of claim 20 to a mammalian subject.

30. A method of inhibiting binding of a ligand to a mlk protein, said method comprising administering a therapeutically effective amount of a composition of claim 20.

31. A method of treating a mlk-related condition, said method comprising administering a therapeutically effective amount of a composition of claim 26 to a mammalian subject.

32. A method of inhibiting binding of a ligand to a mlk protein, said method comprising administering a therapeutically effective amount of a composition of claim 26.

33. A method of identifying an inhibitor of ligand binding to mlk protein which comprises:

(a) combining a mlk protein of claim 14 with a ligand, said combination forming a first binding mixture;

(b) measuring the amount of binding between the mlk protein and the mlk in the first binding mixture;

(c) combining a compound with the mlk protein and the ligand to form a second binding mixture;

(d) measuring the amount of binding in the second binding mixture; and

(e) comparing the amount of binding in the first binding mixture with the amount of binding in the second binding mixture;

wherein the compound is capable of inhibiting ligand binding to mlk protein when a decrease in the amount of binding of the second binding mixture occurs.

34. An inhibitor identified by the method of claim 33.

35. A pharmaceutical composition comprising the inhibitor of claim 34 and a pharmaceutically acceptable carrier.

36. A method of treating a mlk-related condition, said method comprising administering a therapeutically effective amount of a composition of claim 35 to a mammalian subject.

37. A method of inhibiting binding of a ligand to a mlk protein, said method comprising administering a therapeutically effective amount of a composition of claim 35.

38. A composition comprising an antibody which specifically reacts with a ligand of claim 24.

39. A method of promoting bone or cartilage growth, said method comprising administering a therapeutically effective amount of a composition of claim 20 or 26.

40. A method of inhibiting bone or cartilage growth, said method comprising administering a therapeutically effective amount of a composition of claim 20 or 26.

41. A method of inhibiting bone loss, said method comprising administering a therapeutically effective amount of a composition of claim 20 or 26.

FIG. 1

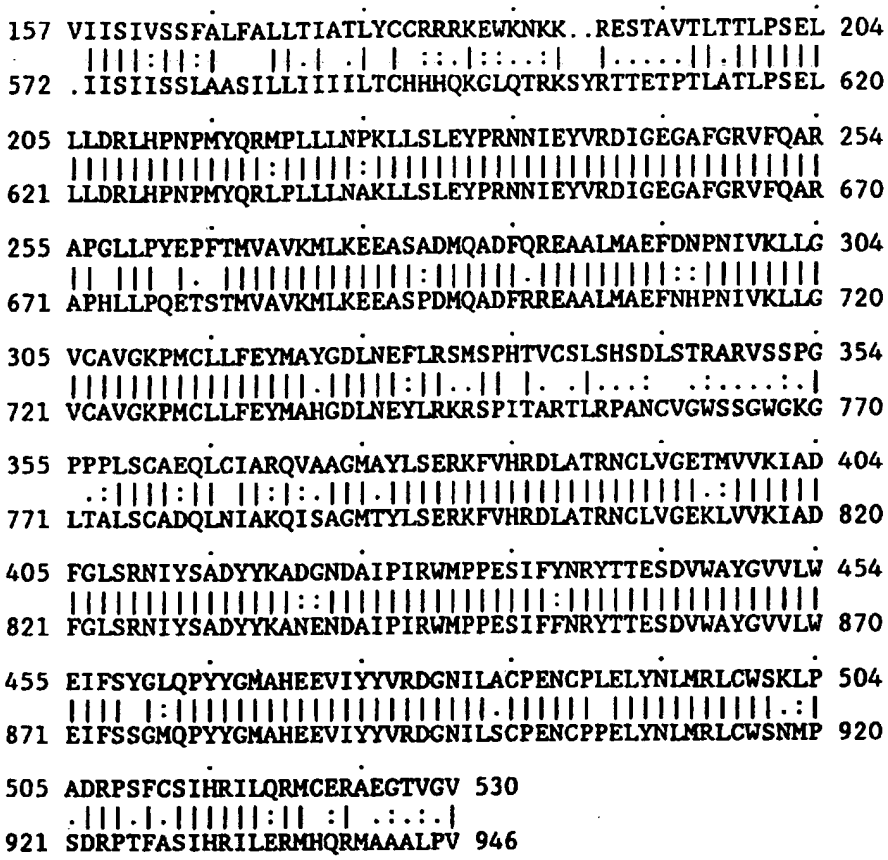

157 VIISIVSSFALFALLTIATLYCCRRRKEWKNKK..RESTAVTLTTLPSEL 204
572 .IISIISSLAASILLIIIILTCHHHQKGLQTRKSYRTTETPTLATLPSEL 620
205 LLDRLHPNPMYQRMPLLLNPKLLSLEYPRNNIEYVRDIGEGAFGRVFQAR 254
621 LLDRLHPNPMYQRLPLLLNAKLLSLEYPRNNIEYVRDIGEGAFGRVFQAR 670
255 APGLLPYEPFTMVAVKMLKEEASADMQADFQREAALMAEFDNPNIVKLLG 304
671 APHLLPQETSTMVAVKMLKEEASPDMQADFRREAALMAEFNHPNIVKLLG 720
305 VCAVGKPMCLLFEYMAYGDLNEFLRSMSPHTVCSLSHSDLSTRARVSSPG 354
721 VCAVGKPMCLLFEYMAHGDLNEYLRKRSPITARTLRPANCVGWSSGWGKG 770
355 PPPLSCAEQLCIARQVAAGMAYLSERKFVHRDLATRNCLVGETMVVKIAD 404
771 LTALSCADQLNIAKQISAGMTYLSERKFVHRDLATRNCLVGEKLVVKIAD 820
405 FGLSRNIYSADYYKADGNDAIPIRWMPPESIFYNRYTTESDVWAYGVVLW 454
821 FGLSRNIYSADYYKANENDAIPIRWMPPESIFFNRYTTESDVWAYGVVLW 870
455 EIFSYGLQPYYGMAHEEVIYYVRDGNILACPENCPLELYNLMRLCWSKLP 504
871 EIFSSGMQPYYGMAHEEVIYYVRDGNILSCPENCPPELYNLMRLCWSNMP 920
505 ADRPSFCSIHRILQRMCERAEGTVGV 530
921 SDRPTFASIHRILERMHQRMAAALPV 946

Fig. 2

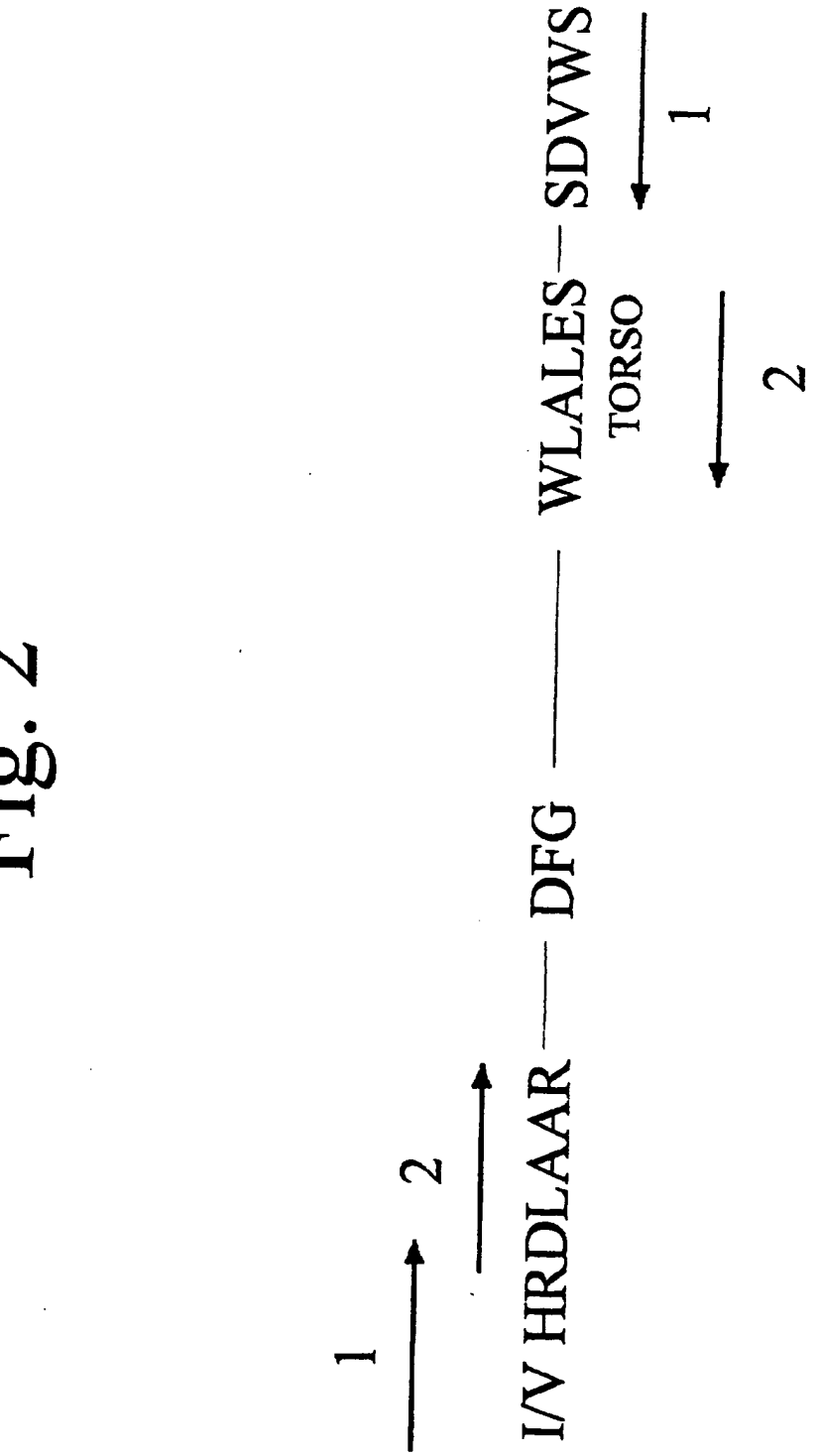
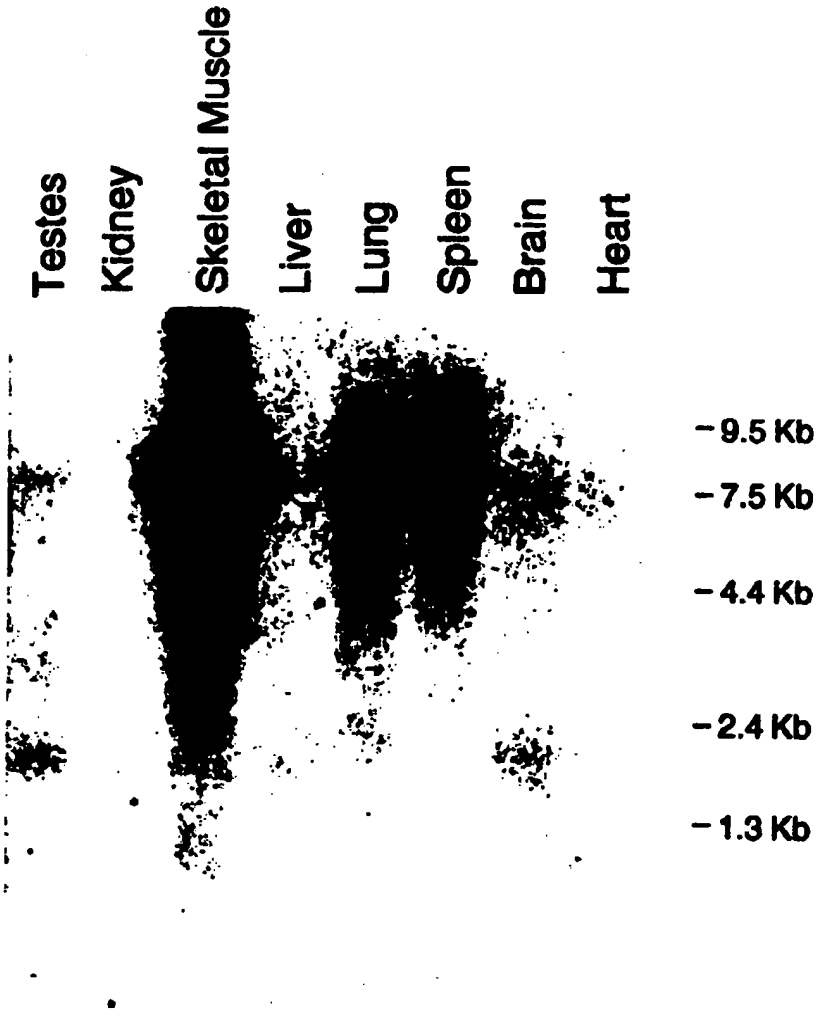


Fig. 3


Testes
Kidney
Skeletal Muscle
Liver
Lung
Spleen
Brain
Heart
9.5 Kb
7.5 Kb
4.4 Kb
2.4 Kb
1.3 Kb

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 95/08493

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 C07K16/18 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 599 274 (SQUIBB BRISTOL MYERS CO) 1 June 1994 see abstract; example 5.7 8 9; table II ---	24-28,38
A	PNAS, vol. 90, 1993 pages 2895-2899, C.G.B.JENNINGS ET AL. 'Muscle-specific trk-related receptor with a kringle domain defines a distint class of receptor tyrosine kinases' see the whole document --- -/--	1-19

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 December 1995	03.01.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Gurdjian, D

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International dication No
PCT/US 95/08493

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NEURON , vol. 10, 1993 pages 963-974, D.M.VALENZUELA ET AL. 'alternative forms of rat trkC with different functional capabilities' see the whole document -----	1-19

1

INTERNATIONAL SEARCH REPORT

Interr nal application No.
PCT/US95/08493

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 29-32,36-37,39-41
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although these claims are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 95/08493

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0599274	01-06-94	AU-B-	5180493	09-06-94
		CA-A-	2103323	25-05-94
		FI-A-	935175	25-05-94
		JP-A-	7132084	23-05-95
		NO-A-	934163	25-05-94